

Three Essays on the Party Strategies under the  
Mixed-Member Electoral System

by

Michio Umeda

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
(Political Science)  
in The University of Michigan  
2011

Doctoral Committee:

Associate Professor Allen D. Hicken, Co-Chair  
Assistant Professor Kenneth Mori McElwain, Co-Chair  
Professor Kenneth W. Kollman  
Professor Walter R. Mebane Jr.

## ACKNOWLEDGEMENTS

My research has tremendously benefitted from the generosity of my dissertation committee. I thank Professors Allen Hicken, Kenneth Mori McElwain, Walter Mebane, and Ken Kollman for their efforts.

I would also like to thank Professors John Campbell, William Clark, Daisuke Nakajima, Anna Grzymala-Busse, Robert Salmond and Rocío Titiunik for their help. For funding fieldwork in Japan, I thank the Center for Japanese Studies and the Horace H. Rackham School of Graduate Studies.

Many colleagues and friends at Michigan have offered very valuable support. In particular, Thomas Chadeaux, Jonson Porteux, and Alton B. H. Worthington deserve special thanks for reading various dissertation chapters and editing my writing. I would also thank to Matias Bargsted, Cassandra Grafström and Su-Hyun Lee for their comments on dissertation chapters.

I thank Professors Kentaro Fukumoto, Yusaku Horiuchi, Gregory Noble, Steven Reed, Jun Saito, Masaki Taniguchi and Naoko Taniguchi for their comments, advice and help to my dissertation research.

Finally, this dissertation is dedicated to my parents, Reiko and Isao Umeda, and my sister, Yukiko, and my brother, Tatsuya.

# TABLE OF CONTENTS

ACKNOWLEDGEMENT	ii
LIST OF TABLES	iv
CHAPTER	
1. Introduction	1
2. Myth of the Contamination Effect?: Different Questions, Different Conclusions for Small Party's Strategy under the Mixed-Member Electoral System	16
3. Moving the Mountain: Strategic Small-Party Nominations to Shift Major-Party Policy Position under Plurality Electoral System	51
4. Challenging Strongmen with Zombies: Incumbency Advantage and the Consequence of the Double Candidacy Claus	86
5. Conclusion	113

## LIST OF TABLES

### TABLE

Table 2.1:	Types of Treatment	48
Table 2.2:	Estimates of Contamination Effect	49
Table 3.1:	SDPJ Nomination in 2005	82
Table 3.2:	DPJ Candidate Position in 2005	83
Table 4.1:	Incumbency Advantage with RDD	111
Table 4.2:	Incumbency Advantage with Natural Experiment	112

# **Chapter 1**

## **Introduction**

In this dissertation, I study the effects of electoral reform on party systems and electoral campaigns, based on the case of Japan following the change to a mixed-member electoral system. The dissertation project examines the issue from three different perspectives: two studies focus on the small party survival strategies under the new system, but from different points of view, and one focuses on the changes in incumbency advantage in plurality contests, following reform.

Electoral systems attract the interest of not only political scientists, but also politicians and citizens. They have significant impact upon the party system, and ultimately upon representation in a democratic system. It is not exaggeration to say that the electoral system is the one of the most influential institutions in a democratic political system, likely second only to the constitution itself.

The electoral system influences democratic representation through various channels: 1) party systems established under the system, 2) policy positions that political parties choose, 3) the policy-making process, given the partisan balance established under the electoral system, and 4) political purposes pursued by the voters at the election.

First, political scientists have argued about the effect of electoral systems on party systems, through the number of parties competing in the electoral districts, since Duverger's seminal work (1959). Duverger argues that plurality electoral systems, like those in the US and Britain, lead to two-party systems, and the proportional representation (PR) systems common in the continental countries, lead to multiparty systems.

There are several studies that test his claim, which is often called “Duverger’s Law”. As early as 1959, Wildavsky discusses that the law works at the district level, but not directly at the national level. Moreover, he also argues that the law regulates the maximum number of the parties in the contest, not the specific number of parties (Wildavsky 1959). In other words, under the plurality electoral system the number of viable candidates could be one, not necessarily two at the district level. Besides, national party system would be single-party dominance or multi-party. More recently, Cox (1997) extends Duverger’s law to a more general “M+1 rule”: no more than M+1 candidates are viable within a district if M is the district magnitude. Moreover, Golder (2006) shows empirical evidence for these arguments: political parties, especially smaller parties, are more likely to form pre-electoral alliances if there are more than two parties competing under a plurality electoral system. Chhibber and Kollman (2004) explain the level of linkage between the district and national level party systems with the centralization of fiscal structure in federal states.

Next, the electoral system influences the policy positions pursued by political parties, and the distribution of policy positions of those represented in the legislature. Downs (1957) argues that with two-party competition under the plurality electoral system, the issue positions of the two parties converge to that of a median voter. In contrast, under a PR system with multiple parties, political parties pursue considerably different policy positions. The outcome depends on many theoretical assumptions of the basic Downsian spatial model (Grofman 2004). If there is a potential entry of third parties to the contest, for example, even under the plurality electoral system the two established parties could have spatially separated positions - more attentive to the preference of the radical constituencies - to deprive small parties of a chance of victory (Palfrey 1984). Moreover, many scholars have reexamined the empirical validity of the Downsian model, especially based on the recent polarization of the US Congress, which consists of two major parties and the members elected with a first-past-the post system (e.g. Iverson 1994; Rabinowitz and Macdonald 1989). However, it is difficult to deny that the electoral system has influence on the variance of the policy positions of parties represented in the legislature.

Moreover, the policy making processes have quite different styles in countries with different electoral systems (Lijphart 1999). Under the plurality electoral system and

subsequent two-party system – the Westminster model -- one party in general achieves majority of the legislature and forms a single-party government. On the other hand, usually no parties achieve legislative majority under a PR system. Therefore, the parties form the coalition governments based on pre- or post-electoral negotiations. Therefore, in the former the single party majority governments run countries in a decisive manner, but in the latter the coalition governments have policy process based on the consensus of the coalition members.

Furthermore, voters elect representatives for different purposes under the different electoral systems. Powell (2000) argues that under the plurality electoral system, the electorates choose a single party in an election that controls the policy making process until the next election. In contrast, under the PR system the voters choose their agents to represent them in a negotiated policy process after the election. Moreover, in the next election the electorates reward or punish the governing party in the former system, which has been delegated decisive power and consequently is accountable for good or bad in the nation during the tenure. In addition, they also decide their voting prospectively based on their judgment regarding which party – the incumbent or the opposition – is more preferable to hold a mandate. Accountability is not so clear in the PR case, because the power is shared by multiple parties in the coalition. Consequently, Powell and Whitten (1993) show that the electorates are more likely to vote based on their evaluation of the government performance, in particular macro-economic situation of the country.

As is clear from the discussion so far, the electoral systems have been broadly represented by two types, single-member plurality and PR systems (Grofman and Lijphart 1986). However, scholars have also recognized that there are a few other systems that achieve some proportional objective with a non-PR system, such as multiple member districts with single non-transferable vote (MMD/SNTV, Lijphart, Pintor and Sone 1986), preferential vote systems (Wright 1986), and mixed-member electoral systems (Lijphart 1986). I focus upon the last in this dissertation project.

There are two sub-groups in the mixed-member electoral systems, based on different linkages between the two electoral tiers. One is the mixed-member proportional system (MMP), currently used in Germany and New Zealand. The system basically

determines the allocation of total legislative seats based on the parties' vote-share in the PR tier, and uses the plurality tier auxiliary to provide priority for the winners in the plurality district among the party candidates. The other is the mixed-member majoritarian (MMM) system that uses two tiers separately. The votes cast for plurality tier do not influence the seat allocation in the PR tier and vice versa (Shugart and Wattenberg 2001).

Historically, only a few countries have used the mixed-member electoral systems. Therefore, study of the mechanisms of the system has been problematic, because of the difficulty in distinguishing the effect of the electoral system from the country-specific factors. For example, (West) Germany is one of the few countries that have been using the system for many years. However, it is not easy to examine the effect of the electoral system on the number of parties, for instance, only based on the Germany case, because not only the electoral system but Germany-specific factors such as social cleavage structure (ex. historical divisions between the liberal and secular Free Democratic Party and the conservative and religious Christian Democratic Union constituencies), or peculiar historical legacy (ex. the Party of Democratic Socialism in the former East Germany area) must have influenced the outcomes that scholars observe.

In the late 1980s and the early 1990s, the mixed-member electoral systems were adopted by both newly democratized countries in the Central and Eastern Europe and countries with long democratic histories, such as Japan, Italy and New Zealand. Following the reforms in these countries, many scholars have examined the function of the mixed-member electoral systems, especially how majoritarian or proportional the systems are, and consequently how many political parties remain and survive under the system (e.g. Cox and Schoppa 2002; Ferrera and Herron 2005; Ferrera, Herron and Nishikawa 2005; Shugart and Wattenberg 2001).

Some studies imply that the mixed-member electoral systems represent the best of both worlds – single-member plurality and PR – because the systems allow both the party-centered campaigning of the PR system and the candidate responsiveness of the plurality system (Shugart and Wattenberg 2001). Nevertheless, the mechanisms of the systems are not yet well understood, even though twenty years have passed since a variety of countries adopted the electoral system.



The MMP and the MMM systems have been regarded as having a similar function, because of their similar appearance that combines the plurality and the PR tiers into an electoral system. In the first few elections following reform, they seem to work in comparable ways. However, it turns out that there are non-ignorable differences between them which are only revealed over time. The MMP system works more like the pure PR system does, but the MMM system works more like the plurality electoral system does, especially if the majority of seats are allocated to the plurality tier.

Initially, a considerable number of small parties ran their candidates in the plurality contests under the MMM system, without a pre-electoral alliance. Some scholars have argued that the MMM system has a considerable proportional tendency, ascribing the strategy change to the effects of the PR tier on the plurality tier: the contamination effect (e.g. Cox and Schoppa 2002; Ferrera, Herron and Nishikawa 2005). The contamination effect is defined as the spillover effect between two electoral tiers in the mixed-member electoral systems, especially the effect of small parties' nominations in plurality contests on the PR tier.

However, Reed shows that the number of parties competing in plurality contests are consistently declining toward two as Duverger/Cox predict (Duverger 1959; Cox 1997) in both Italy and Japan, two countries that adopted the MMM system (Reed 2001; 2007). Maeda also casts doubt on the size of the contamination effect (Maeda 2008), starting from a significant drop in the effective number of electoral parties (Taagapera and Shugart 1989) in plurality contests. These results imply that some adjustment period is required before the electoral reform alters the equilibrium established under the previous system, and consequently the earlier works prematurely rushed to a conclusion before they observed the real outcome of the reform.

In this dissertation project, I examine the effect of adopting MMM system, based on the development of Japanese politics. Therefore, it might be worthwhile to summarize the case since the pre-electoral reform.

In 1994, Japan adopted the MMM electoral system for the House of Representatives (HR). The Liberal Democratic Party (LDP) lost its power in 1993 for the first time since its formation in 1955 because the party lost a considerable number of members due to a party

split prior to the 1993 HR election. After the election, Morihiro Hosokawa, the leader of the Japan New Party (JNP) – a new party formed in 1992 and supported by independents through advocating political reform -- was chosen as Prime Minister of the coalition government, which consisted of eight parties. After long negotiations, the government finally succeeded in passing the electoral reform proposal that changed the electoral rule from the MMD/SNTV to the MMM system.

The MMD/SNTV system had been used since 1947. The electoral system was criticized as a major source of corruption and pork-oriented politics in Japan (e.g. Curtis 1971; Kohno 1997; Ramseyer and Rosenbluth 1993; Reed 1994). Under the system, if any party wants to attain a majority in the legislature, the party has to nominate more than one candidate in a district (which usually elects three to five members for each). Rather than advocating the policy program shared by the other co-partisan candidates, these candidates have incentive to appeal to voters using personal benefits, such as pork-provision and constituent service. In addition, people believed that the generation-long LDP rule lead to corruption and collusion between politicians and bureaucrats. Therefore, the advocates of the reform argued that the reform should change the electoral competition to be more policy oriented, in addition to more frequent turnover.

Based on criticism of the previous MMD/SNTV system, the new electoral system adopted the MMM system that has a single member plurality electoral system consisting of 300 districts as a primary tier. The system also has a PR tier with 200 seats (reduced to 180 in 1998) distributed to 11 regional districts. The reform is intended to have candidates, especially those of the two major parties, compete over policy rather than pork or personal ties. In the single member plurality tier, political parties have no incentive to nominate more than one candidate in each district, and hence the candidates should fight according to the party platform. Moreover, the candidates listed on the party PR list campaign for their parties, for their own survival. The PR tier is also designed as compensation for smaller parties, which have difficulty surviving under in plurality competition. The Japanese voters experienced elections under the current system in 1996, 2000, 2003, 2005 and 2009.

There was a significant party realignment in the 1990s. Therefore, it has been unclear how the electoral reform influenced the party system in Japan. However, the party

system was gradually transformed into a two party system under the new electoral system, consisting of the LDP and the Democratic Party of Japan (DPJ). The LDP returned to the office in 1994, soon after the electoral reform legislation, following the break-up of the JNP coalition. The LDP succeeded in maintaining its power base, in particular by regaining its dominant status in rural areas, and stayed in office as the coalition government until 2009. The DPJ, by contrast, was formed in 1996 by the center-left politicians who were members of the Japanese Socialist Party (JSP) and other small center-left parties as the second largest opposition party at that time. The party became the largest opposition in 1998 following the split-up of the New Frontier Party, and grew to be the unique and serious contender to LDP rule when the Liberal Party merged with the DPJ in 2003 before the HR election. Although the DPJ lost badly to the LDP in 2005, the party made a sweeping victory in 2009 and seized power from the LDP.

In contrast, smaller parties gradually lost support from the electorates, and exited from plurality contests. For example, the Socialist Democratic Party of Japan (SDPJ), formed by the leftist members of the JSP in 1996, campaigned independently until 2005. However, their PR vote-share declined from 8% in 2000 to 6% in 2005. In addition, the number of candidates that the party ran in the plurality contests decreased from 76 in 2000 to 45 in 2005. The party finally decided to form a pre-electoral alliance with the DPJ in 2009, but achieved only 2% of the PR vote-share. Next, another leftist party, the Japanese Communist Party (JCP) also struggled to survive under the system. The JCP had been nominating its candidates to all districts since the MMD/SNTV period, though the party had little chance to get their candidates elected under the plurality system. The party's PR vote-share declined from 11% in 2000 to 7% in 2005, and 5% in 2009. The JCP finally gave up nominating candidates in every district. The party ran candidates in 275 districts out of 300 in 2005, and only 152 districts in 2009. As a result, the average of the effective number of electoral parties (Taagapera and Shugart 1989), which was 2.95 in 1996 and 2.77 in 2000, radically decreased to 2.41 in 2003, to 2.40 in 2005 (Maeda 2008), and further to 2.26 in 2009.

From 2003 to the present (March 2011), only a minor change has happened in the Japanese system consisting of two major parties – the LDP and the DPJ -- and the other smaller parties. Some of the LDP members left to form new parties, such as the People's New

Party in 2005, and Your Party in 2009. Moreover, a religious sect, Happy Science, established a new party with a rightist policy platform. This party, the Happiness Realization Party, ran candidates in most of the plurality districts in 2009. However, these new rightist parties are failing to mobilize significant support from conservative voters. Your Party succeeded in obtaining some support from urban voters in the 2007 House of Councilors (HC) election, although the party's support rate is currently trailing those of the two major parties.

The DPJ's victory in 2009 implies a transformation of electoral campaigns in Japan, because the LDP had been dominant in Japanese politics, thanks to its organizational grip on rural voters (via candidates' personal networks) and public project provision (e.g. Curtis 1971; Saito 2010). Reed, Scheiner and Thies (2009) argue that party affiliation of candidates became a stronger predictor of electoral victory than individual attributes in the 2009 HR election. However, the DPJ does not seem to be immune to personal and/or pork-oriented politics, as evidenced by the scandals of the former party leader Ichiro Ozawa.

The purpose of this dissertation project is to examine the effect of electoral reform on the development of a new party system and party competition. I use the Japanese case, focusing on the evolution of both issues from the 1994 electoral reform to the present. How does a new electoral system regulate the strategy of political parties, and consequently influence the number of parties competing under the system? Did the reform achieve the original goals discussed by advocates? Namely did reform make the electoral campaign more policy oriented and increase electoral turnover?

I investigate these broad questions in specific ways from Chapter 2 to 4. I do not wish to claim that the topics discussed in this work cover all important aspects of the questions regarding the electoral system. However, I believe that these works considerably contribute to our understanding of the electoral system.

First, I study the strategic behavior of small parties in Japan facing a new electoral system that could be biased against them. The project starts with a scholarly interest in the number of parties that keeps competing in the mixed-member electoral systems, particularly in the plurality electoral tier, because it has a substantive impact on the representation mechanism of the system, through the new party system that emerges after the electoral reform.

In Chapter 2, titled “Myth of the Contamination Effect?: Different Questions, Different Conclusions for Small Party’s Strategy under the Mixed-Member Electoral System,” I reexamine accepted wisdom on the contamination effect which posits that it is rational for small parties to nominate candidates under plurality tier of the mixed-member electoral system, because these candidates help to mobilize additional PR votes in their districts (Cox and Schoppa 2002; Ferrera, Herron and Nishikawa 2005; Herron and Nishikawa 2001; Mizusaki and Mori 1998; Reed 2003). Given the effect, smaller parties have more incentives to run candidates in the plurality tier but are less likely to form the pre-electoral alliance, even if their candidates have little chance of victory. Ultimately, the effect could influence the representation process by shifting the optimal strategies of smaller parties under the mixed-member electoral system.

In the chapter, I demonstrate that the mixed results of previous studies are due to different assumptions about information available to political parties, but not to researchers, based on Rubin’s Causal Model (RCM, Holland 1986). By using the genetic matching approach (Diamond and Sekhon 2005), I show that small parties in fact increased their PR vote-share where they nominated candidates in plurality tier. Nevertheless, I conclude that the impact was not large enough at the national level for small parties to stay in the plurality contests, because the average effect size would have been much smaller if they had increased the number of districts in which they nominated, compared to those in which they actually did. However, this outcome leaves the original question unanswered. Why do small parties nominate candidates in plurality tier, despite their low likelihood of victory and the marginal gain in PR vote-share from nomination?

In Chapter 3 titled “Moving the Mountain: Strategic Small-Party Nominations to Shift Major-Party Policy Position under Plurality Electoral System,” I show alternative purpose of nominations: blackmailing major parties into adopting policies more desirable to the small party. This chapter explains why small parties often run their candidates in plurality electoral systems in general, not necessarily in the plurality tier of a mixed-member electoral system. Moreover, it also describes why major parties in plurality electoral systems often choose less centrist positions, despite the Downsian theory that predicts they would be better off by choosing a position closer to the median voter.

I develop a formal model of the strategic interactions between a small leftist party and the center-left major party which extends the chain-store paradox (Kreps and Wilson 1982). The model shows that the small party can persuade the major party to select leftist, rather than centrist, candidates in a repeated game, regardless of the median voter's preference towards the centrist candidates. Since the small party threatens to run its own leftist candidate, despite the risk of benefiting the rightist party, the major party shifts the policy position. The empirical data from the Japanese case clearly supports the model: the small leftist SDPJ ran candidates if the center-left/major DPJ nominated centrist candidates, and in turn the DPJ selected leftists if the SDPJ was popular in the districts, because the nominations could be a serious threat to their competition against the rightist LDP.

Next, in Chapter 4 titled “Challenging Strongmen with Zombies: Incumbency Advantage and the Consequence of the Double Candidacy Claus,” I study the magnitude of incumbency advantage in Japan after a transition to the mixed-member system. The incumbency advantage – the electoral advantage that the incumbent candidates maintain, thanks to their status as the incumbent -- could be harmful to democratic accountability because it helps both unpopular incumbent candidates and unpopular governing parties (with more incumbent MPs) retain power.

As mentioned above, Japanese elections have been described as “candidate-centered” since the MMD/SNTV era. The candidate-centered campaign is expected to create a large incumbency advantage, because the electorate focuses more on pork and narrowly-focused interests provided by powerful incumbent politicians utilizing the resource available from their office, rather than the policy platforms presented by political parties. Despite the reputation for candidate-centered campaigns, however, I show that Japanese incumbents lack this advantage. To demonstrate this, I utilize Regression Discontinuity Design, which has recently been applied in political science to estimate incumbency advantage in plurality electoral systems (e.g. Ariga 2010; Hainmueller and Kern 2008; Lee 2008; Linden 2004).

By using a natural experiment, I demonstrate that the double-candidacy system explains this finding. I utilize a mechanism of this system that provides PR seats for candidates in a regional PR district who lost in plurality contests, but not for those in other

PR districts with very similar electoral results. Because the regions do not have a direct effect on the electoral outcome, the difference of the electoral outcome between these districts in the next election is attributable to the advantage of the PR-incumbent candidate. The design shows that these candidates have a 2-3% advantage from their incumbency of the PR tier.

Counter to the conventional wisdom (e.g. McKean and Scheiner 2000), this result infers that the double-candidacy system hurts local incumbents by awarding available PR seats to the most competitive “losing” candidates (at the district level) in each party. In the next election, these MPs challenge at the district level with incumbency status from their PR seats, cancelling out the advantage of district incumbency. In other words, the incumbent MPs have benefitted from their incumbency status, as the traditional view of Japanese campaigning implies. However, the benefit is often counter-balanced by the benefit to their opponents from PR incumbency, especially in the most competitive districts. This result explains why the LDP failed to keep its dominant status, particularly in rural Japan.

Finally, Chapter 5 concludes the dissertation with a summary of the findings, in addition the directions of the future in political science demonstrated with the work.

## REFERENCE

- Ariga, Kenichi. 2010. "Entrenched Incumbents, Irresponsible Parties? Comparative Analysis of Incumbency Advantage across Different Electoral Systems." Dissertation manuscript
- Chhibber, Pradeep K., and Ken Kollman. 2004. *The Formation of National Party Systems: Federalism and Party Competition in Canada, Great Britain, India, and the United States*. Princeton, N.J.: Princeton University Press.
- Cox, Gary W. 1997. *Making Votes Count: Strategic Coordination in the World's Electoral Systems*. Cambridge: Cambridge University Press.
- Cox, Karen E., and Leonard J. Schoppa. 2002. "Interaction Effects in Mixed-Member Electoral Systems: Theory and Evidence from Germany, Japan, and Italy." *Comparative Political Studies* 35 (9): 1027–1053.
- Curtis, Gerald L. 1971. *Election Campaigning, Japanese Style*. New York: Columbia University Press.
- Diamond, Alexis and Jasjeet S. Sekhon. 2005. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." Working Paper.  
<http://sekhon.berkeley.edu/papers/GenMatch.pdf>
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper & Row.
- Duverger, Maurice. 1959. *Political Parties, Their Organization and Activity in the Modern State*. Array London: Methuen.
- Ferrara, Federico, and Erik S. Herron. 2005. "Going It Alone? Strategic Entry under Mixed Electoral Rules." *American Journal of Political Science* 49 (1):16–31.
- Ferrara, Frederic, Erik S. Herron, and Misa Nishikawa. 2005. *Mixed Electoral Systems: Contamination and Its Consequences*. Palgrave Macmillan, New York. N.Y.
- Golder, Sona Nadenichek. 2006. "Pre-Electoral Coalition Formation in Parliamentary Democracies." *British Journal of Political Science* 36 (2): 193-212.
- Grofman, Bernard. 2004. "Downs and Two-Party Convergence." *Annual Review of Political Science* 7: 25-46.
- Grofman, Bernard and Arend Lijphart ed. 1986. *Electoral Laws and Their Political Consequences*. New York: Agathon Press.



- Hainmueller, Jens and Holger L. Kern. 2008. "Incumbency as a Source of Contamination in Mixed Electoral Systems." *Electoral Studies* 27 (2): 213-227
- Herron, Erik S. and Misa Nishikawa. 2001. "Contamination effects and the Number of Parties in Mixed- Superposition Electoral Systems." *Electoral Studies* 20(1): 63-86.
- Holland, Paul. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81(396): 945-970.
- Iverson, Torben. 1994. "The Logics of Electoral Politics: Spatial, Directional, and Mobilizational Effects." *Comparative Political Studies* 27(2): 155-189.
- Kohno, Masaru. 1997. *Japan's Postwar Party Politics*. Princeton, N.J.: Princeton University Press.
- Kreps, David M, and Robert Wilson. 1982. "Reputation and Imperfect Information." *Journal of Economic Theory* 27 (2): 253-279.
- Lee, David S. 2008. "Randomized Experiments from Non-Random Selection in US House Elections." *Journal of Econometrics* 142 (2): 675-697.
- Lijphart, Arend. 1986. "Proportionality by Non-NR Methods: Ethnic Representation in Belgium, Cyprus, Lebanon, New Zealand, West Germany and Zimbabwe". In Bernard Grofman and Arend Lijphart ed. *Electoral Laws and Their Political Consequences*. New York: Agathon Press. 113-123
- Lijphart, Arend. 1999. *Patterns of Democracy: Government Forms and Performance in Thirty-six Countries*. New Haven: Yale University Press.
- Lijphart, Arend, Rafael Lopez Pintor, and Yasunori Sone. 1986. "The Limited Vote and the Single Nontransferable Vote: Lessons from the Japanese and Spanish Examples." In Bernard Grofman and Arend Lijphart ed. *Electoral Laws and Their Political Consequences*. New York: Agathon Press. 154-169
- Linden, Leigh L. 2004. "Are Incumbents Really Advantaged? The Preference for Non-Incumbents in Indian National Elections." Unpublished manuscript.
- Maeda, Ko. 2008. "Re-examining the Contamination Effect of Japan's Mixed Electoral System Using the Treatment-Effects Model." *Electoral Studies* 27(4): 723-731
- McKean, Margaret and Ethan Scheiner. 2000. "Japan's New Electoral System: la plus ça change..." *Electoral Studies* 19 (4): 447-477.

- Mizusaki, Tokifumi and Hiroki Mori. 1998. "Tokuhyou deeta kara mita heiritsusei no mekanizumu." *Senkyo Kenkyu [Electoral Studies]* 13: 50-59.
- Palfrey, Thomas R. 1984. "Spatial Equilibrium with Entry." *The Review of Economic Studies* 51(1): 139-156.
- Powell, G. Bingham, Jr. 2000. *Elections as Instruments of Democracy: Majoritarian and Proportional Visions*. New Haven: Yale University Press.
- Powell, G. Bingham and Guy D. Whitten. 1993. "A Cross-National Analysis of Economic Voting: Taking Account of the Political Context." *American Journal of Political Science* 37: 391-414.
- Ramseyer, J. Mark, and Frances McCall Rosenbluth. 1993. *Japan's Political Marketplace*. Cambridge, Mass.: Harvard University Press.
- Rabinowitz, George, and Stuart Elaine Macdonald. 1989. "A Directional Theory of Issue Voting." *American Political Science Review* 83 (1): 93-121
- Reed, Steven R. 1994. "Democracy and the Personal Vote: A Cautionary Tale from Japan." *Electoral Studies* 13(1): 17-28
- Reed, Steven R. 2003. "Heiritsusei ni okeru shousenkyoku kouhosha no hireidaihyou tokuhyouritsu eno eikyou [The Effect of the PR Tier on SMD Nomination Strategies in Mixed Member Systems: Some Evidence from Japan]." *Senkyo Kenkyu [Electoral Studies]* 18: 5-11.
- Reed, Steven R., Ethan Scheiner, and Michael Thies. 2009. "New Ball Game in Politics." *Oriental Economist* (October 2009): 8-9
- Saito, Jun. 2010. *Jiminto Choki Seiken no Seiji Keizaigaku [The Political Economy of the LDP Regime]*. Tokyo: Keiso Shobo.
- Shugart, Matthew S., and Martin P. Wattenberg ed. 2001. *Mixed-member Electoral Systems: the Best of Both Worlds?* Oxford, UK: Oxford University Press.
- Taagapera, Rein and Matthew S. Shugart. 1989. *Seats and Votes: The Effects and Determinants of Electoral Systems*. New Haven: Yale University Press
- Wildavsky, Aaron B. 1959. "A Methodological Critique of Duverger's Political Parties" *The Journal of Politics*, 21(2): 303-318

Wright, Jack F.H. 1986. "Australian Experience with Majority-Preferential and Quota-Preferential Systems." In Bernard Grofman and Arend Lijphart ed. *Electoral Laws and Their Political Consequences*. New York: Agathon Press. 124-138

## Chapter 2

### **Myth of the Contamination Effect?: Different Questions, Different Conclusions for Small Party's Strategy under the Mixed-Member Electoral System**

#### INTRODUCTION

This chapter discusses the *contamination effect* in the mixed-member electoral system and the methodological issues regarding the estimation of the effect size. More specifically, I argue the problems that arise when actors decide the treatment assignment according to information not necessarily observable to researchers – *selection bias* (Angrist 2004) and *essential heterogeneity* (Heckman, Urzua and Vytlačil 2006) -- based on the recent development of statistics in regards to causal inference from the observational data. The mixed-member electoral system uses two traditional electoral systems, the single-member plurality (SMP) and the proportional representation (PR) systems as the two electoral tiers of the system. The *contamination effect* is defined as the spillover effect of the small parties' nominations in plurality contests to the PR tier as their additional vote-shares. Previous studies find contrasting estimates of the effect. This chapter shows that their differences are due to inconsistencies between the specific questions posited and the methodological tool used, in addition to potential problems of their data and models of these studies. The argument of this chapter is not only applicable to the discussion regarding the *contamination effect*. It is also useful for any political studies that attempt to estimate the effect of the *treatment*, which actors select the units to receive the treatment based on the information not necessarily shared with scholars.

The *contamination effect* is attracting the interest of many political scientists because of its potential effect to the electoral strategies of small parties under the mixed-member electoral system (Cox and Schoppa, 2002; Ferrera, Herron and Nishikawa, 2005; Herron and Nishikawa, 2001; Mizusaki and Mori, 1998; Maeda 2008; Reed, 2003). If the small parties could mobilize a considerable amount of additional PR votes by nominating their candidates in the plurality tier, they would be less motivated to form pre-electoral alliances with other parties in the plurality contests or to merge with other parties for their survival. Hence, the effect could change how majoritarian the electoral system is, and consequently transform the party system as well.

In the 1990's, many democratic countries adopted the mixed-member electoral system. However, not only politicians but also scholars had little knowledge of the mechanism due to the historical rareness of the system. In these countries, scholars witnessed that a substantial number of small parties kept their candidates in plurality contests, even though these candidates had a small chance of victory. The *contamination effect* is expected to explain their strategies: these small parties nominate their candidates in plurality contests because they help the parties to mobilize more PR votes.

Previous studies show inconsistent results for the *contamination effect*. Several studies observe the positive and statistically significant effect of contamination. (Cox and Schoppa 2002; Ferrera, Herron and Nishikawa 2005; Herron and Nishikawa 2001; Mizusaki and Mori 1998; Reed 2003) In contrast, recently Maeda (2008) criticizes that the estimates of those studies are biased due to *selection bias* because the parties select the districts to nominate in which they are stronger. In other words, he assumes the pre-campaign party popularity of the districts is not balanced between those the parties nominate and those they do not, even if we control all the covariates observable to scholars. He argues that the *contamination effect* is not significant once the bias is corrected with the Heckman's treatment-effects model (Maddala 1983).

I agree with Maeda that the estimates of the previous studies could have been biased. Nevertheless, I do not think that the problem is serious. The source of *selection bias* in the estimation of the *contamination effect* is the information gap between the

parties that choose the districts to nominate and the scholars who attempt to analyze the effect of the nomination, regarding the party popularity in each district before the campaign. The countries that adopted the mixed-member electoral system have experienced considerable number of elections under the system, and the party systems in those countries become stable along with the new electoral rule. Consequently, scholars now have more reliable sources of the pre-campaign party popularity in each district, although it might not match perfectly with those held by the parties. Moreover, in this chapter I demonstrate that Maeda's model specification is justifiable only under the assumption that the *contamination effect* does not exist in the previous election.

On the other hand, in this chapter I argue that it is more reasonable to assume party discretion regarding the heterogeneity of the treatment effect, or *essential heterogeneity* (Heckman, Urzua and Vytlačil 2006): the parties nominate to districts in which they can mobilize more PR votes than in the other districts. (For example, due to different quality of activists prepared to run in each district.) The potential heterogeneity of the treatment effect could be an important barrier to attaining reasonable inference in regards to the *contamination effect*. However, the previous studies overlook this point based on the implicit assumption of the conventional statistical approach of treatment effect *homogeneity*.

If the parties choose the districts to assign their candidates according to the knowledge of the heterogeneous treatment effect, there are two theoretically relevant *contamination effects*. One is *how much the party actually gained in the districts where the party nominated candidates*, which is termed as *the average treatment effect for the treated (ATT)*. The other is *how much the party would have gained if the party randomly chose the districts to nominate*, termed as *the average treatment effect (ATE)* (Rosenbaum and Rubin 1983).

Based on these theoretical discussions, this chapter analyzes the *contamination effect* with the Japanese case, which Maeda also uses. The results clearly support the *contamination effect* in the both definitions. Moreover, the results show that the *ATT* is larger than the *ATE*: the parties have at least partial information regarding the different size of the treatment effect among the districts, and they exploit the information to

choose the districts to achieve the best outcome with the minimum cost. This result suggests potential problem in using estimates from previous studies to infer a counterfactual case for assessing the potential impact of the *contamination* effect on the party system.

The chapter proceeds as follows: The next section describes the theoretical background and the previous studies regarding the *contamination effect*. Then, I demonstrate the difference in the theoretical concepts to be measured by respective approaches. In the empirical section, I analyze the Japanese case for which the previous studies reached different conclusions. After describing the empirical background of the cases, the subsequent section reports the outcomes of the analysis. The results clearly support the *contamination effect*. The final section discusses both the empirical and theoretical implications of this chapter.

## **THE CONTAMINATION EFFECT: HOW THE MIXED-MEMBER ELECTORAL SYSTEM WORKS?**

Since the seminal work of Duverger (Duverger 1959), the electoral system has attracted the interest of several political scientists due to its effects on the party system, and ultimately its effect on representation in the democratic system. Duverger argues that plurality electoral systems like those in the US or the UK lead to two-party systems, and PR electoral systems lead to multiparty systems (Duverger 1959). There are several studies following his claim: Duverger's law. As early as 1959, Wildavsky argues that the *law* primarily works at the district level but not directly at the national level. Moreover, he discusses that it regulates the maximum number of the parties in contests, not the number of parties (Wildavsky 1959). More recently, Cox extends the Duverger's law in more general "M+1 rule": no more than M+1 candidates are viable in the district level if M is the district magnitude (Cox 1997). Furthermore, Chibber and Kollman discuss the linkage between the district and national level party system with the centralization of budget in the federal states (Chibber and Kollman 2004).

As Duverger explains, electoral systems are broadly represented with the single-member plurality (SMP) and the proportional representation (PR) systems (Grofman and Lijphart 1986), although scholars have recognized that there are a few other systems that achieves proportional representation with the non-PR system (Lijphart 1986). On the other hand, in the 1990s both newly democratized countries in the Central and Eastern Europe and the countries with long democratic histories such as Japan, Italy and New Zealand adopted the mixed-member electoral system. Scholars are interested in the system because of its hybrid nature that combines two traditional electoral systems, the SMP and the PR systems into the single electoral system<sup>1</sup>.

Due to its historical rareness, it had been hard to study the mechanisms of the system because of the difficulty to distinguish the effect of the electoral system from the country specific factors. For example, although Germany has been using the system for many years, it is hard to differentiate the effect of the electoral system on the number of electoral parties from Germany-specific factors, such as social cleavages (ex. historical division between the liberal and secular SPD and the conservative and religious CDU/CSU constituencies) or peculiar historical legacy (ex. the PDS in the former DDR area). Some studies suggest that the mixed-member electoral system represents the best of both worlds (Shugart and Wattenberg 2001), because the system allows both party-centered campaigning of the PR system and candidate responsiveness of the plurality system. Nevertheless, the mechanisms of the system are not well understood yet, because only few countries have used it for a significant period of time in the past.

Initially, some scholars attempted to use the two electoral tiers of the mixed-member electoral system as the natural laboratory to test the theory regarding the electoral system such as the Duverger's law, because different electoral systems are

---

<sup>1</sup> There are two sub-categories of the system based on different linkage of the two electoral tiers. One is the mixed-member *majoritarian* system (MMM) adopted in Japan that uses two tiers separately. The other is the mixed-member *proportional* system (MMP) used in Germany and New Zealand, which basically determines the allocation of total legislative seats according to the party's vote-share in the PR tier, and uses the plurality tier auxiliary to give priority for the winners among the party candidate (Shugart and Wattenberg 2001).



treated to the parties and the electorates within homogeneous social and historical environments (Moser and Scheiner 2004). An old, but still important, criticism of the theory of electoral systems is the endogeneity of the electoral system to the social structure that precedes modern nations, such as the number of social, ethnic, or religious cleavages within the society. Hence, rather than the electoral system influencing the party system, the number of the social cleavage regulates the party system, and in turn the party system decides the electoral system where the parties compete (Cox 1997). The mixed-member electoral system is expected to be a solution to the endogeneity because two electoral tiers could be used for the *controlled comparison*, at which all the covariates are held constant between two treatment groups.

However, subsequent studies criticize the idea of the *controlled comparison*, because of the interactions of two electoral components (Cox and Schoppa 2002; Ferrera, Herron and Nishikawa 2005). Ferrara and his colleagues argue that the results of each electoral component are contaminated from the other electoral tier, because the political parties and the electorates do not make strategic decisions separately as if the two components were different elections. The parties design the nomination strategy to maximize the total number of the seats from two tiers. Moreover, the electorate is simultaneously exposed to the campaign messages tailored for both components. Hence, the electoral outcome of each electoral tier should not be used for the *controlled comparison*, because the outcomes are contaminated with the influence of the other component.

These studies demonstrate two pieces of evidence of contamination. One is the higher vote-share that parties attain in PR tier where the parties run their candidates at plurality tier (Cox and Schoppa 2002; Ferrera, Herron and Nishikawa 2005; Herron and Nishikawa 2001; Mizusaki and Mori 1998; Reed 2003). This is the *contamination effect* in a narrow definition, which is the main topic of this work. Given the effect, smaller parties might have incentives to run their own candidates in the plurality tier even if they have little chance of victory rather than to form the electoral alliance with other parties or merge into larger parties, because they contribute to mobilize more PR votes. Hence, the

effect encourages small parties to campaign alone despite the electoral disadvantage for smaller parties in majoritarian electoral systems (Duverger 1959; Cox 1997). As a result, the number of parties that compete in the plurality tier of the mixed-member electoral system will be inflated – the second evidence of the *contamination effect* that the scholars show. Cox and Schoppa show that the number of the parties in Japan measured with effective number of electoral parties (Taagapera and Shugart 1989) was 2.95 in 1996 and 2.77 in 2000, which was much larger than those in the pure plurality system: the average is 2.24 among them (Cox and Schoppa 2002).

However, there are mutually interconnected problems for those alleged evidence of the *contamination effect*. First, inflated numbers of the candidates in a plurality tier cited as a result of the *contamination effect* could be a mere consequence of the temporal coordination problem among the parties under the unstable party system. The subsequent elections suggest the temporal nature of the inflation. As Duverger's law implies, the district-level competition quickly moved closer to a two party system in Italy and Japan, two countries with the mixed-member majoritarian electoral system (Reed 2001; 2007). Moreover, the number of parties that compete in plurality districts decreased rapidly as the parties adjust their strategies to survive under the system, such as to form the electoral alliance or to merge with the other parties. In Japan, the average effective number of electoral parties at the district-level went down to 2.41 in 2003 and 2.40 in 2005 (Maeda 2008), and to 2.26 in the 2009 election. This number is almost the same with the countries with pure plurality electoral system.

Moreover, there are two potential problems with using the estimates of the *contamination effect* in the previous studies to infer the impact of its effect on the party system. One is *selection bias*: if scholars have less information than political parties regarding the pre-campaign party support level in each district (*pretreatment*), the estimate of the *contamination effect* might be biased due to the correlation between the nomination and the *pretreatment* party's popularity in these districts, which is observable to the parties but not to the scholars.

For scholars, it is difficult to have a reliable measurement for pre-campaign support level of each party in newly demarcated districts for the first few elections after the electoral reform. To estimate the effect, some earlier studies compare the PR vote-share between the districts with and without party candidates without any control of the pre-campaign support level (Mizusaki and Mori 1998). The others use some indirect indicators of the party popularity such as the existence of the incumbent candidates (Cox and Schoppa 2002; Ferrera, Herron and Nishikawa 2005). These studies make valuable contribution as a preliminary analysis of the effect. Nevertheless, their result is vulnerable to *selection bias* due to the insufficient control of the pre-campaign party support, and hence should be revised once more reliable covariates become available to scholars.

Maeda recently criticized these estimates of the earlier scholarship as being biased by the *selection bias* (Maeda 2008). In other words, he claims that parties have the private information regarding the pre-campaign support level in each district, which is unobservable to researchers, and utilize the information to select the districts to nominate. From the Japanese case, Maeda argues that the *contamination effect* is insignificant once the *selection bias* is corrected with the treatment-effects model (Maddala 1983).

However, there are two problems in Maeda's conclusion. One is the model specification: his specification is almost equivalent to claim that *the contamination effect does not exist given the assumption that it did not exist in the previous election*. The treatment-effects model requires that the covariates that have direct effect to the outcome variable be used for the outcome function. Nevertheless, he does not include the nomination in the last election to the covariates of the outcome function. The specification is justifiable only if the *contamination effect* did not exist in the previous election.

The other is the magnitude of the information gap between the parties and scholars. It is true that the estimation of the *contamination effect* is vulnerable to *selection bias* if scholars employ weak indicator of the pre-campaign party support in each district

such as the incumbency. However, as we experience more elections after the electoral reform, scholars attain more reliable information source in regards to the pre-campaign party popularity in each district, such as the party's PR vote-share in the last election. Hence, the information disparity between scholars and parties regarding the pre-campaign party support, which could be the source of the *selection bias* if the party exploits, would not be large even if it might exist, given the current availability of the information to scholars. In the following section, I indicate that the treatment-effects model is more supportive to the *contamination effect* if the specification problem is corrected.

Nevertheless, there is another problem to use these estimates for assessing the impact of the *contamination effect*, for example its potential impact on the national party system. Previous discussion of the *contamination effect* argue the rationality to have independent campaign for small parties in plurality tier of the mixed-member electoral system instead of forming the pre-electoral alliance or the merger with the other parties, given the disadvantage for small parties under the plurality contests. Therefore, they often use their estimates of the *contamination effect* to infer the impact of the effect what if the parties nominate as many candidates as possible, in almost all districts, for instance.

However, these discussions implicitly rely on the *homogeneity* assumption of the effect size across districts. In other words, they assume that the effect size would have been the same if the parties had nominated the candidates to the districts that they actually did not. This assumption, is widely used in most of political science studies without any doubt or knowledge, but is not necessarily true, especially if the actors could choose the units that receive the treatment. For example, political parties might nominate in the districts where the parties could achieve larger *contamination effect* by using their private information regarding the quality of local party activists prepared to run for the campaign. Consequently, if these parties had nominated to the districts that they did not, they would not have attained as much average gain as they actually did. This *heterogeneity* of the treatment effect – called as *essential heterogeneity* – has recently gained the attention of several scholars (ex. Heckman, Urzua and Vytlačil 2006). The

strength of the *essential heterogeneity* assumption is that the parties have strong incentive to gain and utilize the information regarding the effect size because they can skim the benefit from the information.

If the treatment effect is not constant, there are two theoretically relevant definitions of the *contamination effect*. One is *how much the party actually gained in the districts where the party nominated candidates*, which is termed as *the average treatment effect for the treated (ATT)*. The other is *how much the party would have gained if the party randomly chose the districts to nominate*, termed as *the average treatment effect (ATE)*. If the question is whether the *contamination effect* exists or not, the *ATT* could be more appropriate parameter given that the parties have no obligation to randomly choose the districts to nominate. On the other hand, if scholars attempt to evaluate the potential impact of the *contamination effect* if the parties run as many candidates as possible, in other words, much more than they actually did, the *ATE* could be more appropriate parameter because the average effect size would converge to the *ATE* as the number of the districts with the candidates increases to all the districts. In the next section, I indicate the theoretical basis of these different questions and the corresponding statistical methods for the estimation.

## METHOD

This section discusses the theoretical background of *selection bias* and *essential heterogeneity* and the solutions for these problems based on the framework of the *Rubin Causal Model (RCM)* (Holland 1986). According to the convention of the *RCM*, I describe the context of the *contamination effect* as an analogy of the experiment: the *units* of the study are plurality districts, the *treatment* is the nomination of the candidate to the district, and the *outcome* is the PR vote-share of the party inside of the area demarcated by the plurality districts.

Suppose that there are two treatment states (treatment/control), as a party runs a candidate or not in a district.  $Y_1$  is the outcome if the unit receives the treatment and  $Y_0$  is the outcome if the unit receives the control.  $Y_1$  and  $Y_0$  are the function of observed

covariates  $X$  and unobserved covariates  $U$ . Hence, the electoral outcome of a party in a district can be a function of observable covariates  $X$  such as the previous electoral outcome or industrial structure of the district, in addition not easily observable covariates  $U$  such as quality of potential candidates in the district, given their treatment status: the nomination of the candidate.

$$Y_1 = \mu_1(X) + U_1: \text{the outcome } Y \text{ if the unit receives the treatment, given } X \text{ and } U_1 \quad (1-1)$$

$$Y_0 = \mu_0(X) + U_0: \text{the outcome } Y \text{ if the unit receives the control, given } X \text{ and } U_0 \quad (1-2)$$

In the RCM, the causal effect is defined as the difference between two outcomes, treated ( $Y_1$ ) and controlled ( $Y_0$ ). For instance, each district has two potential outcomes for the party A's PR vote-share – the outcome realized if the party A runs a candidate and the outcome attained if the party does not, given the other party's strategy and political environment in the districts. Keeping all the observable covariates  $X$  implicit, the outcome  $Y$  we observe is:

$$\begin{aligned} Y &= D Y_1 + (1 - D) Y_0 \\ &= Y_0 + (Y_1 - Y_0) D \\ &= \mu_0 + (\mu_1 - \mu_0 + U_1 - U_0) D + U_0 \end{aligned}$$

Using the conventional regression notation:

$$= \alpha + \beta D + \varepsilon \text{ such that } \alpha = \mu_0, \beta = \mu_1 - \mu_0 + U_1 - U_0, \text{ and } \varepsilon = U_0 \quad (1-3)$$

However, in reality we can observe only one of the two potential outcomes for each unit. If a unit is treated, we do not observe the outcome of the unit if it were not treated. To estimate the treatment effect from the observational studies, RCM shows that three conditions are required: 1) *stable unit treatment variable assumptions (SUTVA)*, 2) *unconfoundedness*, and 3) *common support*. Rosenbaum and Rubin (1983b) term *unconfoundedness* and *common support* as *strong ignorability*.

The *SUTVA* is lack of the interaction among the units. More specifically, suppose the situation with  $N$  units indexed by  $i = 1, 2, \dots, N$ ;  $D$  types of the treatment is indexed by  $d = 1, 2, \dots, D$ ; and the outcome variable  $Y$ , which possible values are denoted by  $Y_{di}$ . The *SUTVA* is the assumption that the value of  $Y$  for unit  $i$  such that  $\forall i \in N$ , when exposed to the treatment  $d$  will be the same no matter what mechanism is used to assign the

treatment  $d$  to unit  $i$ , and no matter what treatments the other units received (Rubin 1986). Next, *unconfoundedness* is the assumption that the distribution of the outcome  $Y$  is conditionally independent of the treatment  $d$  given the observed covariates  $X$ , i.e.  $(Y_1, Y_0) \perp D \mid X$ . In other words, *unconfoundedness* indicates that there are no unobservable covariates  $U$  that have influence on the outcome  $Y$ , given the observed covariates  $X$ . Finally, *common support* is the condition that the units with the same value of the covariates  $X$  have a positive probability of being the treated or the controlled, hence  $0 < P(D = 1 \mid X) < 1$ . This assumption is required to avoid the extrapolation problem.

Heckman and his colleagues show that instead of *unconfoundedness*, *mean independence* conditional on the observed covariates  $X$ ,  $E(U_0 \mid X, D=1) = E(U_0 \mid X, D=0)$  is sufficient to attain unbiased estimate of the *average* treatment effect (Heckman, Ichimura and Todd 1998). *Mean independence* is a much weaker assumption than *unconfoundedness* because it requires that only the *mean* of unobserved covariates  $U_0$  are balanced between the treated and the controlled groups. In the term of this chapter, *mean independence* conditional on observed covariates  $X$  is equivalent to the absence of *selection bias*. If the *mean independence* is violated – if there is *selection bias* – the statistical model that corrects the bias is required to have unbiased estimate of the treatment effect, such as Heckman's treatment-effects model.

In the following section, I discuss two potential problems to assess the impact of the *contamination effect*, *selection bias* and *essential heterogeneity*, and the corresponding approaches to solve the problems. First, I discuss *selection bias* and Heckman's treatment effect model, to describe a specification problem in Maeda's approach. In addition, the problem could have been important in the earliest studies of the *contamination effect* but now it is expected not to be serious. Next, I describe *essential heterogeneity*, which could be a more important problem in evaluating the consequence of the *contamination effect* from the estimates based on the conventional approach. I alternatively offer the matching approach to estimate the *ATT* and the *ATE*, two relevant parameters to assess the impact of the *contamination effect*.

### Selection Bias and Heckman's treatment effect model

As described, if there is *selection bias* – violation of *mean independence* – the estimation of the treatment effect will be biased without using specific methods to correct the bias. Suppose that the parties know their pre-campaign support level in each district in terms of the attributes not observable to researchers ( $\varepsilon = U_0$ ), and nominate in the districts where the party could have done well without the nomination ( $E(U_0|D=1) > 0$ ): the parties nominate to the districts where they are originally stronger. Under the assumption, the observed difference between the treated  $Y_1$  and the controlled  $Y_0$  given the observed covariates  $X$  is:

$$\begin{aligned} E(Y_1 | X, U, D=1) - E(Y_0 | X, U, D=0) \\ = \beta + E(U_0 | X, D=1) - E(U_0 | X, D=0) \end{aligned} \quad (2-1)$$

If the *mean independence* assumption is satisfied, the term  $E(U_0|D=1) - E(U_0|D=0)$  is cancelled out, and hence the observed difference between the treated and the controlled given  $X$  indicates the treatment effect  $\beta$ . However, if there is *selection bias*, the term is not equal to zero because  $E(U_0|D=1)$  is expected to be positive, and  $E(U_0|D=0)$  is expected to be negative. The observed difference given  $X$  will be the biased estimate of the treatment effect  $\beta$ .

Heckman's treatment-effects model is designed to solve the problem. Suppose that the treatment decision/nomination  $D$  is decided by the observed  $Z$  and the unobservable  $U_D$  as (2-2) shows. We observe that the party runs candidates to the districts ( $D=1$ ) if the underlying parameter  $D^*$ , which is the function of  $Z$  and  $U_D$  exceeds 0 but choose not to participate ( $D=0$ ) if  $D^*$  is below 0, as (2-3) denotes.  $Z$  is assumed to be not correlated with  $\varepsilon$ , and  $I$  is the indicator function. The treatment effect is assumed as constant, therefore  $\eta = U_1 - U_0 = 0$ , and  $\beta = \beta\_bar$ :

$$\begin{aligned} D^* &= \mu d(z) + U_D \\ &= z\gamma + U_D \end{aligned} \quad : \text{Treatment Function} \quad (2-2)$$

$$D = I(D^* > 0) = I(U_D > -\mu d(z)) \quad : I \text{ is an indicator function} \quad (2-3)$$

$$Y = \alpha + \beta D + \varepsilon \quad : \text{Outcome Function} \quad (2-4)$$



The treatment-effects model assumes the joint multivariate normal distributions of the error terms of the treatment and the outcome functions, i.e.  $\varepsilon$  and  $U_D$ .  $\phi(\cdot)$  and  $\Phi(\cdot)$  are the probability and cumulative density functions of the standard normal distributions, and  $\rho = \text{Corr}(\varepsilon, U_D)$  and  $\sigma_\varepsilon = \text{st.dev.}(\varepsilon)$ .  $U_D$  is not directly observable. However,  $U_D$  is expected to be large if the unit is selected into the treatment group even if the observed characteristics of the unit  $Z$  suggest otherwise. In contrast,  $U_D$  is expected to be small if the unit does not join into the treatment group although observed attributes of the unit indicate that the unit is likely to participate. Given the distributional assumptions of the error terms, the observed outcomes of the treated and the controlled are expressed as:

$$\begin{aligned} E(Y_1 \mid X, Z, U, D=1): \text{observed outcomes of the treated} \\ &= \mu_1 + E(\varepsilon \mid U_D > -z\gamma) \\ &= \mu_1 + \rho \sigma_\varepsilon [\phi(z\gamma) / \Phi(z\gamma)] \end{aligned} \quad (2-5a)$$

$$\begin{aligned} E(Y_0 \mid X, Z, U, D=0): \text{observed outcomes of the controlled} \\ &= \mu_0 + E(\varepsilon \mid U_D < -z\gamma) \\ &= \mu_0 + \rho \sigma_\varepsilon [-\phi(z\gamma) / \{1-\Phi(z\gamma)\}] \end{aligned} \quad (2-5b)$$

Therefore, the observed difference between the treated and the controlled given the observed covariates, which is shown in (2-1), is rewritten by the difference (2-5a) and (2-5b):

$$\begin{aligned} (3-1) &= \mu_1 + \rho \sigma_\varepsilon [\phi(z\gamma) / \Phi(z\gamma)] - \mu_0 - \rho \sigma_\varepsilon [-\phi(z\gamma) / \{1-\Phi(z\gamma)\}] \\ &= (\mu_1 - \mu_0) + \rho \sigma_\varepsilon [\phi(z\gamma) / \Phi(z\gamma)(1 - \Phi(z\gamma))] \end{aligned} \quad (2-6)$$

such that  $\beta = (\mu_1 - \mu_0)$ , and  $E(U_0 \mid D=1) - E(U_0 \mid D=0) = \rho \sigma_\varepsilon [\phi(z\gamma) / \Phi(z\gamma)(1 - \Phi(z\gamma))]$ .

Hence, the conventional regression notation is rewritten as

$$\begin{aligned} Y &= \alpha + \beta D + \varepsilon \\ &= \mu_0 + (\mu_1 - \mu_0) D + \rho \sigma_\varepsilon [\phi(z\gamma) / \Phi(z\gamma)] D + \rho \sigma_\varepsilon [\phi(z\gamma) / \Phi(z\gamma)(1 - \Phi(z\gamma))] (1-D) \\ &= \mu_0 + (\mu_1 - \mu_0) D + \rho \sigma_\varepsilon [\phi(z\gamma) (D - \Phi(z\gamma))] / [\Phi(z\gamma) (1 - \Phi(z\gamma))] \end{aligned} \quad (2-7)$$

The treatment-effects model uses *the inverse mills ratio*,  $h(z) = [\phi(z\gamma) (D - \Phi(z\gamma))] / [\Phi(z\gamma) (1 - \Phi(z\gamma))]$  (Heckman 1976) – also termed as *hazard variable* (Barnaw, Cain

and Goldberger 1980) -- to remove the *selection bias*. By adding  $h(z)$  to the conventional OLS model, which is estimated with probit analysis for (2-3), the treatment-effects model is supposed to eliminate the bias, and consequently estimate the unbiased estimate of the treatment effect,  $(\mu_1 - \mu_0)$ .

For the specification of the model, some of the covariates  $Z$  in the treatment function should be also included as the covariates in the outcome function  $X$ , if the variables have direct effect on both the treatment assignment and the outcome. However,  $Z$  should contain at least one covariate not included in  $X$ . The covariates  $Z$  that are not included in  $X$  must satisfy the *exclusion restriction*: these variables should not have any effect on the outcome except through the treatment. Otherwise there will be the correlation between  $Z$  and error term in the outcome function  $\varepsilon$ . If any variable with the direct effect on the outcome is included only to the treatment function, the estimate of the treatment effect will be biased.

This point has significant importance for this chapter. In his treatment-effects model, Maeda uses the nomination to the districts in the last election i.e. the  $t-1$  nomination as the covariates of the treatment function. However, he does not use the  $t-1$  nomination as the covariate of the outcome function. Nevertheless, if the *contamination effect* existed in the last election, the  $t-1$  nomination should have direct -- expected to be negative -- effect on the outcome, because the  $t-1$  electoral outcome in the outcome function should be already *contaminated* by the party candidate. His model specification is justifiable only under the assumption that the *contamination effect* does not exist in the last election.

On the other hand, the significance of the *selection bias* relies on the reliability of the indicator of the pre-campaign support level available to scholars, i.e.  $X$ . The more scholars have good indicators of the support, the smaller the term  $E(U_0|X, D = 1) - E(U_0|X, D = 0)$  will be. In the estimation of the *contamination effect*, the parties' PR vote-share of the last election in each district, which is not available for the earliest studies but now available, could be strong indicator of the party support just before the campaign. With the information, *selection bias* becomes serious problem only if the

following four conditions are satisfied: 1) the shift of the party popularity between elections varies across plurality districts, 2) the party has private information regarding the heterogeneity in the shift, 3) the party decides nomination based on the information, and 4) other covariates that scholars use in their model failed to explain the shift. In other words, the party should nominate not in the districts that the party is strong but those that the party becomes stronger since the last election, and scholars do not know where the party becomes stronger. I am rather skeptical that these conditions are satisfied.

#### Essential Heterogeneity and Matching Approach

This section discusses *essential heterogeneity* – the heterogeneity of the treatment effect across the units. Although most of political studies implicitly or unknowingly assume the homogeneity of the effect size, *essential heterogeneity* has significant importance to assess the impact of the treatment, especially if the actors in the study have discretion to choose the units to receive the treatment, and/or scholars attempt to answer a counterfactual question. For example, “How large would the average treatment effect be if the actors significantly increase the number of units to receive the treatment?” The discussion in regards to the *contamination effect* is a good example that *essential heterogeneity* matters. Political parties are expected to have at least partial information regarding heterogeneity, such as the quality of the local activists prepared to run. Moreover, previous studies often attempt to evaluate the potential impact of the effect – for example, parties nominating as many candidates as possible -- based on the estimate of the effect for a relatively small number of the districts in which political parties actually nominate their candidates.

The *average treatment effect*, the ATE is defined by the following formula.

$$\begin{aligned}
 ATE &= E(Y_1 - Y_0 | X, U) \\
 &= E(Y_1 | X, U) - E(Y_0 | X, U): \text{given additive separability} \\
 &= \mu_1 - \mu_0
 \end{aligned}
 \tag{3-1}$$

Practically, given *SUTVA*, *mean independence* and *common support*, the *ATE* could be estimated by integrating  $X$  for its range. :

$$(3-1) = E(Y_1 | X) - E(Y_0 | X): \text{given mean independence for } U \text{ given } X$$

$$\begin{aligned} &= \int_{x \in X} [E(Y_1 | X) - E(Y_0 | X)] dF_X \\ &= \int_{x \in X} [E(Y_1 | X = x) - E(Y_0 | X = x)] P(X = x) dF_X \end{aligned} \quad (3-2)$$

On the other hand, the treatment effect for the units that actually received the treatment is called as the *ATT*, *the average treatment effect for the treated*. The *ATT* is the difference in the outcome between the observed outcomes for the treated units – the PR vote-share of the party A in the districts that the party nominated candidates -- and potential outcomes for the units what if they did not receive the treatment – the PR vote-share of the party A what if the party had decided not to nominate. If the actors deciding the treatment assignment have even partial information regarding the heterogeneity of the treatment effect, there could be a sorting on the gain:  $Cov(D, \beta) > 0$ , and consequently  $Cov(D, \eta) > 0$  because  $\mu_1 - \mu_0$  is constant. In this case, the *ATT* should be larger than the *ATE* because  $Cov(D, \eta) > 0$ .

The *ATT* is denoted with the formula as follows:

$$\begin{aligned} ATT &= E(Y_1 - Y_0 | X, U, D=1) \\ &= E(Y_1 | X, U, D=1) - E(Y_0 | X, U, D=1): \text{given additive separability} \end{aligned} \quad (3-3)$$

The estimation of *ATT* requires weaker assumptions than those for the *ATE*. The *mean independence* is required only for the controlled:  $E(U_0 | X, D=0) = 0$ , and the *common support* is required only for one side:  $P(D = 1 | X) < 1$  (Heckman, Ichimura and Todd 1998). If these weaker assumptions are satisfied, the *ATT* is estimated with:

$$(3-3) = E(Y_1 | X, D=1) - E(Y_0 | X, D=1): \text{given mean independence for } U_0 \text{ given } X \text{ and } D=1.$$

$$\begin{aligned} &= \int_{x \in X | D=1} [E(Y_1 | X) - E(Y_0 | X)] dF(X | D=1) \\ &= \int_{x \in X | D=1} [E(Y_1 | X = x) - E(Y_0 | X = x)] P(X = x | D=1) dF(X | D=1) \end{aligned} \quad (3-4)$$

The difference between the *ATT* and the *ATE* is that the *ATE* is estimated by measuring the difference in the outcome variable  $Y$  of the treatment and controlled for each value of the covariates  $X$  and weighting the difference according to the distributions of  $X$ , but the *ATT* is estimated by weighting the difference based on the distributions of  $X$  for the treated group ( $D=1$ ).

In this chapter I use the genetic matching approach (Sekhon, Forthcoming) to estimate the *ATT*, in addition to the *ATE* for comparison. The algorithm is an extension of the Mahalanobis metric (Rubin 1980) with the generalized Mahalanobis metric<sup>2</sup>, which includes an additional weight matrix to find the best match between the treated and the controlled. To estimate the *ATT*, the method finds the match between each treated unit and  $M$  most similar controlled units, set to one in this chapter. To estimate the *ATE*, the method also finds the match between the controlled units and  $M$  most similar treated units, in addition to the procedure for the *ATT*.

Moreover, I conduct sensitivity analysis to examine how robust the estimates based on the matching are to the unobserved confounders. Estimates of the treatment effect based on matching are unbiased if there are no unobserved confounders. A sensitivity analysis is designed to provide a quantifiable increase in uncertainty if this key assumption of no unobserved confounders is relaxed (Keele 2009). Rosenbaum (2002) has developed theoretical framework of the sensitivity analysis to measure the robustness of the estimates based on the matching to possible presence of an unobserved confounder. In this chapter, I use the *rbounds* package of *R* (Keele 2009) to the estimates of treatment effect by the matching approach, and indicate the minimum value of  $\Gamma$ : the log odds of differential assignment to treatment due to an unobserved confounder, required to have p-value of the estimates higher than .05.

## EMPIRICAL ANALYSIS

In this chapter, I use the electoral data of the Japanese House of Representatives (HR) to estimate the size of *contamination effects*. Japan adopted the mixed-member majoritarian

---

<sup>2</sup> The metric is indicated in Appendix.

(MMM) electoral system to the House of the Representatives in 1994. The electorates in Japan experienced five elections under the system (1996, 2000, 2003, 2005 and 2009). I focus on two leftist parties in Japan, the *Social Democratic Party of Japan* (the SDPJ in the following) in the 2003 and 2005 elections, and the *Japanese Communist Party* (the JCP in the following) in the 2009 election. I do not analyze the effect for the two major parties, the *Liberal Democratic Party* (the LDP in the following), the *Democratic Party of Japan* (the DPJ in the following), which Maeda analyzed in his chapter, in addition the other elections for the two small leftist parties. There are following theoretical and technical reasons for the choice of parties and elections.

- 1) The availability of the party support indicator of the previous election in the districts
- 2) The fraction of the districts that the parties nominate
- 3) The (lack of) pre-electoral alliance with the other parties

First, as I have discussed, a reliability of the estimate significantly depends on the quality of the information of the party support level in each district before the campaign: *pretreatment*. The parties' PR vote-share of the last election is the most widely used indicator of the analysis. However, this indicator requires that the "same" parties continue to exist between two elections. The Japanese party system has been very fluid between 1994 and 2003, especially around the 1996 election. Hence, for parties that experienced the party realignment before the elections that we analyze, it is difficult to use the PR vote-share in the last election as the indicator of the *pretreatment*.

For example, the DPJ newly nominated candidates in some of the districts in the 2000 and the 2003 elections not because the party recruited new candidates in those districts but because the incumbent candidates of dissolved parties -- the *New Frontier Party* for 2000 and the *Liberal Party* for 2003 -- who switched their affiliation to the DPJ. In these districts, the supporters of these incumbent candidates turned to the DPJ because the DPJ became the party of their favorite candidates, not because the DPJ specifically ran candidates in their districts. My preliminary analysis shows that the DPJ substantively increased the PR vote-share where the party nominated the former

incumbents of the dissolved party, according to the previous vote-share of the dissolved party in the districts<sup>3</sup>. Therefore, the *contamination effect* of the DPJ would be biased without careful attention to avoid the overestimation for these elections.

The JCP and the SDPJ did not experience such a significant party realignment after 1996. The JCP was completely insulated from the party realignment in the 1990s; as well the SDPJ have been almost intact after the 1996 election.

The SDPJ case might need some additional description. The *Japanese Socialist Party* (JSP) had been the largest opposition party during the Cold War period, but the party renamed itself to the SDPJ in 1996 to renew the image. However, the SDPJ lost most of the moderate members and traditional support base such as the labor unions to the DPJ within the year. Moreover, some leftist members deserted the SDPJ to form the (New) JSP in the year, criticizing the compromise that the party leaders made with the coalition partner at that time, the LDP. The New JSP gained some support from the remaining leftist constituents in the 1996 election, though the party lost the eligibility to compete in the PR tier after the election. Hence, in the 2000 election many of the previous voters to the New JSP are expected to vote for the SDPJ. Moreover, the SDPJ failed to participate in one of the regional PR districts in 1996, Hokkaido, owing to the rapid breakup of the party organization during the period especially in the region. Consequently, for the SDPJ the estimate of the contamination effect in 2000 might be less reliable. The leftist constituents who voted for the New JSP in 1996 could have voted for the SDPJ in 2000. However, the analysis uses the electoral outcome of the SDPJ in 1996 as the pretreatment, although the indicator could be biased according to different support level of the New JSP in the election. Conversely, the SDPJ did not experience large party realignment after 1996. Hence, in this chapter I exclude the SDPJ in 2000 from the analysis, which requires that the SDPJ electoral outcome of each district being comparable between 1996 and 2000.

Next, smaller fractions of the districts that the SDPJ and the JCP nominate contributes attaining a reliable estimate, especially that of the *ATT*, because it is not only

---

<sup>3</sup> The table is not included but available from the author by request.

easier to find the good *match* between the treated and the controlled units but also the common support assumption,  $0 < P(D=1|X) < 1$  for the *ATE*, and  $P(D=1|X) < 1$  for the *ATT* is more easily satisfied. Suppose that a party *A* nominates in total  $i$  districts among  $n$  districts. To estimate the *ATT* of the *contamination effect*, the matching approach finds a district from the pool of  $n-i$  districts without candidates, which has the most similar covariate distribution with the district with the candidate, the district of the interest, and ascribes the difference of the outcomes between these two districts to the treatment effect. Therefore, the smaller the ratio  $i/(n-i)$  is, the more similar districts are likely to be picked up to match the districts with candidates, because the district of the reference is chosen from the larger pool of districts in the controlled group.

In the five elections after the electoral reform, the SDPJ nominated no more than one sixth of the total plurality districts. This small fraction, however, is convenient to attain more reliable estimate of the *ATT*. In contrast, the LDP and the DPJ nominated most of the districts for all elections. This makes it difficult to attain good estimate, not only that of the *ATT* but also of the *ATE*, because it is not easy to find the matched units between the treated and the controlled if the number of the controlled units are very small. The JCP on the other hand nominated in all districts until 2003 and 275 districts in 2005, but reduced the number of districts to nominate to 152 in the 2009 election. Therefore, I analyze the *contamination effect* of the JCP in 2009, but not for the other elections.

Finally, the electoral alliance disturbs estimating the genuine effect of the nomination on the PR vote-share. Previous studies on Japan do not analyze the *contamination effect* of another small party, Komei party, because the party has been in the coalition with the LDP since 1997 and fought the campaign under the pre-electoral accord with the LDP: the party nominates in the plurality tier only under the accord with the LDP, and as a result the districts with Komei candidates are different from those without the candidates not only in the existence of the candidate. For example, the LDP publicly requested its supporters to cast their PR votes for Komei where Komei does not run in plurality tier and support the LDP candidates instead. This could result



in the underestimation of the *contamination effect* for the party. Similarly, in the 2009 election the SDPJ formed the pre-electoral alliance with the DPJ, and nominated only in the districts without the DPJ candidates except for few districts. Therefore, I do not to use the electoral outcome in 2009 for the SDPJ. Consequently, in this chapter I use the SDPJ in 2003 and 2005, and the JCP in 2009 for the analysis.

On the other hand, there is another source of the bias in estimation the *contamination effect*. If available, previous studies often use the party's PR vote-share at the district in the previous election as the indicator of the pre-campaign party support level, i.e. *pretreatment*. However, it is problematic to use the variable as the proxy of the *pretreatment*. The PR vote-share could be already contaminated by the presence of the candidates in the last election<sup>4</sup>. We might not observe relative increases of PR vote-share in the district  $i$  between time  $t$  and  $t+1$  even if the party nominates candidate at  $t+1$ , because the party already had the candidate at time  $t$  and the electoral outcome was already influenced by the presence of the candidate.

I use two approaches to solve the problem. First, I add the nomination of the party candidates in the last election as a covariate. The model is termed as the *pooled* model, by contrast with the following separated models. On the other hand, there could be different mechanisms and the size of the *contamination effect* in the districts that the party nominated in the previous election and that it did not. To cope with the potential difference of the districts, I divide the districts into two groups according to whether the districts had the party candidate in the previous election or not, and conduct the analyses separately. For the districts with candidates in the last election, the treatment is *keep*: the party continuously runs candidates in both elections, and the controlled condition is *exit*: the party ran in the last election but does not in the current one. In contrast, for the districts that the party did not nominate in the last election, the

---

<sup>4</sup> Reed mentions this point and estimates the *contamination effect* separately between *keep* and *exit* on the one hand, and *participate* and *keep out* on the other hand (Reed 2003).

treatment is *participate*: the party newly participates to the plurality contest in the current election, and the controlled is *keep out*: the party does not participate in both elections.

Table 2.1 indicates the number of districts that belong to each category for the SDPJ and the JCP. There are 242 districts eligible for the analysis of the 2003 election<sup>5</sup>. Among these districts, the SDPJ nominated in 54 districts in the 2000 election. In the 2003 election, the party kept nominating candidates in 29 districts out of 54 districts, in addition to newly participating in 19. Hence, the party nominated candidates in total at 48 districts among the districts available to the analysis<sup>6</sup>. In short, in the 2003 election the SDPJ had 29 *keep*, 25 *exit*, 19 *participate* and 169 *keep out*. In the 2005 election, 297 districts are available for the analysis. From 61 districts with the SDPJ candidates in 2003<sup>7</sup>, the party selects 26 districts to keep nominating candidates in 2005, in addition participated in 12 districts. Therefore, the SDPJ had candidates in 38 districts, as well had 26 *keep*, 35 *exit*, 12 *participate* and 224 *keep out* districts in the 2005 campaign.

In contrast, the JCP nominated at 272 districts out of 297 available districts in 2005. Among the 25 districts that the party did not nominate in 2005, the JCP newly nominate only in a district in 2009. Hence, the party has only one district of *participate*. Therefore, for the analysis of the JCP in 2009, I analyze only the effect of *keep* (150 districts) against *exit* (122 districts).

I use three covariates for the matching to balance the *pretreatment* – the pre-campaign party support level in the districts. First, I use the PR vote-share of the last election as a main covariate. In addition, I also employ the urbanization level of the

---

<sup>5</sup> Owing to the redistricting in 2002, 57 districts are not available for 2003 because these districts contain different geographic areas between elections. In addition, a district (Kanagawa 14th) is not available for the analysis of 2003 because of the reorganization of the municipalities. Similarly, three districts are not available for the analysis in 2005 because of the reorganization of the municipalities in Ibaraki prefecture. Consequently, out of 300 districts, 242 for 2003 and 297 for 2005 are available for the analysis.

<sup>6</sup> Out of 300 total districts, the SDPJ nominated in 62 districts (Maeda 2008).

<sup>7</sup> The number includes districts not eligible to the analysis of the 2003 election.

districts measured by the DID index<sup>8</sup>, which has been widely used in the analyses of Japanese election. If the party's relative support level did not change between two elections, matching on the previous PR vote-share balances the party support level before the campaign. Moreover, even if the party's support level has shifted between the elections, and the party has the information regarding the shift to choose the districts to nominate, the potential imbalance of the *pretreatment* due to the information is considerably reduced if the shift is significantly correlated with the urbanization level of the district. Finally, only for the pooled model, I use the existence of the party candidate in the last election as a covariate, because the variable is constant in the separated models.

In addition, to compare the results with the methods in the previous studies, I estimate the *contamination effect* with the pooled data of the SDPJ by the *OLS* with the same covariates with the matching model in this chapter. Moreover, I also estimate the effect by using the treatment-effects model with the same variables that Maeda uses, except using the nomination in the previous election to the outcome function to correct his specification problem.

## RESULT

Table 2.2 reports the results of the analysis in regards to the 2003 and 2005 elections for the SDPJ and that of 2009 for the JCP. The table includes the estimates from the conventional *OLS* and the treatment-effects model for the SDPJ, to compare the outcomes with those in the methods of the previous studies.

In 2003, the SDPJ attained on average 2.1% more votes in the districts that the party nominated the candidates to the plurality contests than the districts that the party did not. Similarly, the party also gained on average 2.1% and 2.2% more PR votes where the party kept the candidates or newly participated than it exited or kept out. The

---

<sup>8</sup> The urbanization level of the district is measured by Densely Inhabited District (*DID*) index, the concept of the Japanese census that indicates the fraction of the population living in the densely populated area.

estimations of the *ATT* are very stable among different conditions. On the other hand, the sizes of the effects are relatively heterogeneous in 2005. The party on average attained 2.7% additional PR votes in the districts that the party nominated in the pooled data. However, the party gained 3.7% additional votes where the party kept the candidates than where it exited, but it achieved less than 1% of gain -- statistically insignificant, at the districts that it newly participated. All covariates are relatively well balanced between the treated and the controlled units.

The estimates of the *ATE* by the matching approach indicate the smaller effects: 1.9% in the 2003 pooled data, and 1.3% in the 2005 pooled data. The numbers should be used carefully because the matching approach failed to achieve good balance of the covariates between the treated and the controlled, because of small number of the districts with the candidates<sup>9</sup>. However, the difference between the estimates of the *ATT* and the *ATE* suggests that the discretion of the parties to choose the districts matter. The SDPJ did not randomly choose the districts to nominate candidates, but instead carefully choose the districts to bring better PR outcome to the party.

The conventional *OLS* models with the same covariates estimate the effect as 1.7% and 1.9% for the pooled data in 2003 and 2005<sup>10</sup>. The treatment-effects models that use the nomination in the previous election as the covariates of the outcome function, estimate the *contamination effect* as 1.6% and significant in 2003, and 1.1% and insignificant in 2005. The estimates are now much similar to those of the *OLS* or the *ATE* estimates than the original estimates by Maeda's model, which indicate even negative effect of the *contamination effect*.

---

<sup>9</sup> This failure of the balancing suggests the potential extrapolation problem behind the conventional approaches (ex. King and Zeng 2006; King and Zeng 2007).

<sup>10</sup> Yitzhaki (1996) demonstrates that the weighting scheme of the *OLS* regression is influenced to the distance to the median of the distribution, as well the distance to the adjacent observations. Heckman and Vytlačil (2005) also indicate that the *OLS* weight scheme is special because the *OLS* weights for the generalized Roy model do not necessarily integrate to 1 nor are they not necessarily non negative. See Yitzhaki (1996) and Heckman and Vytlačil (2005) for the details of their discussions.

Finally, the result of the JCP shows that the *contamination effect* is not only the phenomenon of the SDPJ. The JCP, which conducts least candidate-centered campaign in Japanese politics, also achieved additional PR votes by nominating candidates to the plurality contests. The JCP on average attained .88% additional PR vote-shares in the districts that the party kept nominated in 2009. The size of the *contamination effect* is much smaller than those of the SDPJ in the previous campaigns, but the effect is still statistically significant.

The result of sensitivity analysis indicates that the estimates based on the matching are very robust to the potential presence of an unobserved confounder. There should be very significant unobserved confounders: 4.0 to 6.0 of  $\Gamma$  for the *ATT* estimates to fail rejecting the null hypothesis at the conventional  $p=.05$  level. It is difficult to imagine such a large imbalance exists between the treated and the controlled. If the party has decided the nomination based on the level of party support in each district unobservable to researchers, as Maeda claims, the party must have very accurate information for the support level given the large values of  $\Gamma$ .

## DISCUSSION AND CONCLUSION

The results of the analyses indicate that both the SDPJ and the JCP surely attained extra PR vote-shares in the districts where the parties nominated candidates, except in those where the SDPJ participated for the first time in 2005. The size of the *contamination effect* is in general even larger than those estimated with the conventional *OLS*. For the SDPJ, the effect was 2-4% in the districts that the party actually nominated. This is quite substantive for a small party like the SDPJ that attained on average 4-6% of the PR vote-share in those elections. Hence, the case of the SDPJ in 2003 and 2005 supports the claim that the small parties are in fact able to mobilize additional PR vote by nominating candidates in plurality contests, contrary to Maeda's doubts as to the existence of the *contamination effect*. The size of the effect is much smaller for the JCP. It could be the result of the less candidate-based campaign of the JCP, in addition to the potential for a

large number of less qualified JCP candidates. Despite this however, the JCP were still able to mobilize additional PR support in those districts by the nomination.

Nevertheless, the implication of the results could be more supportive to Maeda in a substantive as opposed to statistical significance in the context of Japanese politics<sup>11</sup>. First, the size of the effect is not small but probably not enough to motivate small parties to keep candidates in the plurality contests without electoral accord with other parties. The estimates of the *ATE* suggests that the SDPJ would have gained on average 1-2% more PR votes if the party randomly chose the districts to run their candidates, though there is some risk of bias due to a covariate imbalance. Suppose that the party had nominated to all the districts, and the size of the treatment effect does not change even if the party distributes organizational resources thinly across the districts, as the *SUTVA* supposes. Even under these strong assumptions, the party would have gained at most an additional seat from some regional PR lists with relatively large district magnitude, such as Kinki with 29 PR seats (3.45% PR share/seats) or South Kanto with 22 (4.55% PR share/seats).

Finally, in addition to these *quantitative* problems, some clauses of the Japanese electoral law raise the hurdle for smaller parties to nominate for as many plurality districts as possible in order to utilize the *contamination effect*. Although they received little attention in the previous studies, the Japanese electoral law imposes a financial penalty for minor candidates that fail to achieve 10% of the valid ballots of the district<sup>12</sup>. These clauses charge extra costs for small parties to nominate additional candidates, in addition to the regular costs of recruitment and campaign activities.

---

<sup>11</sup> Reed (2003) has mentioned the substantive impact of the effect beyond “statistical significance” of the coefficient. He argues that the size of the *contamination effect* is at most to provide a couple of additional seats for small parties, but not so substantive for them under the mixed member majoritarian system.

<sup>12</sup> In the Japanese HR election, candidates are required to deposit 3,000,000 yen (33,000 US\$) when they register to run for the plurality districts. The deposit is confiscated if they fail to achieve 10% of the valid ballot. (*Public Office Election Law, Article 92 and 93*). Moreover, those candidates lose the eligibility for public funding to some of their electoral expenses (*ibid. Article 141 Clause 7, Article 142 Clause 10, Article 143 Clause 14 and Article 164-2 Clause 6*).

In short, the analysis of this chapter supports the existence of the *contamination effect* itself. Nevertheless, the effect would not be sufficient to have a small party keep their candidates in the plurality district contests, at least to maximize their seats with a reasonable cost. The result does not preclude the possibility that small parties continue running in the plurality tier to achieve other purposes, such as to make a credible threat to other major parties so that they adopt more desirable policies for the small parties, which is discussed in the next chapter of this dissertation. The parties are also able to nominate candidates under the electoral accord with other parties, as the SDPJ did in the 2009 election with the DPJ in most of the districts. However, the result of the chapter implies that it was no accident that small parties gradually exited from the plurality contests in Japan, and as a consequence the effective number of parties went down to the level of the countries with the pure plurality electoral system.

As a more general implication, this chapter demonstrates the importance of elaborating the concept of the study and behavioral assumptions of the actor that decide the treatment assignment based on the theoretical framework of the statistics. I demonstrate the importance of the assumptions regarding the information available to the actor but not to researchers, in addition to the incentives of the actors to exploit the information. Moreover, I also describe the difference in the concepts: the average treatment effect for the districts where the party actually nominated candidates (*ATT*), and the average treatment effect if randomly chosen districts are treated (*ATE*). The elaboration of these concepts is necessary to choose the appropriate analytical tool for the question. The theory of statistics regarding the causal inference made tremendous developments in the last decade. Some of the developments are already applied in the field of political science. This work demonstrates another example of fruitful direction of future research, which explores unresolved questions in political science with a more refined approach.

## REFERENCE

- Barnow, B.G., Glen G. Cain and Arthur S. Goldberger. 1980. "Issues in the Analysis of Selectivity Bias." *Evaluation Studies* 5: 43-59.
- Chibber, Pradeep and Kenneth Kollman. 2004. *The Formation of National Party Systems: Federalism and Party Competition in Canada, Great Britain, India, and the United States*. Princeton, N.J.: Princeton University Press.
- Cox, Gary W. 1997. *Making Votes Count: Strategic Coordination in the World's Electoral Systems*. Cambridge: Cambridge University Press.
- Cox, Karen E. and Leonard J. Schoppa. 2002. "Interaction Effects in Mixed-Member Electoral Systems: Theory and Evidence from Germany, Japan, and Italy." *Comparative Political Studies* 35(9): 1027–1053.
- Duverger, Maurice. 1959. *Political Parties, Their Organization and Activity in the Modern State*. Array London: Methuen.
- Ferrara, Federico and Erik S. Herron. 2005. "Going It Alone? Strategic Entry Under Mixed Electoral Rules." *American Journal of Political Science* 49 (1): 16–31.
- Ferrara, Frederic, Erik S. Herron and Misa Nishikawa. 2005. *Mixed Electoral Systems: Contamination and Its Consequences*. Palgrave Macmillan, New York.
- Grofman , Bernard and Arend Lijphart ed. 1986. *Electoral Laws and Their Political Consequences*. New York: Agathon Press.
- Heckman, James J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models." *Annals of Economic and Social Measurement* 5, 475–492.
- Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1): 153-61.
- Heckman, James J., Hidehiko Ichimura, Jeffrey Smith and Petra Todd. 1998. "Characterizing Selection Bias Using Experimental Data." *Econometrica* 66(5): 1017-1098.
- Heckman, James J. and Salvador Navarro-Lozano. 2004. "Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models." *Review of Economics and Statistics* 86(1): 30-57.



- Heckman, James J., and Richard Robb. 1985. "Alternative Methods for Evaluating the Impact of Interventions: An Overview." *Journal of Econometrics* 30(1-2): 239-67.
- Heckman, James J., Justin L. Tobias and Edward J. Vytlačil. 2001. "Four Parameters of Interest in the Evaluation of Social Programs." *Southern Economic Journal* 68(2): 210-223.
- Heckman, James. J., Sergio Urzua and Edward J. Vytlačil. 2006. "Understanding Instrumental Variables in Models with Essential Heterogeneity." *Review of Economics and Statistics* 88 (3): 389-432.
- Heckman James J. and Edward J. Vytlačil. 2005. "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica* 73(3): 669-738.
- Herron, Erik S. and Misa Nishikawa. 2001. "Contamination effects and the Number of Parties in Mixed- Superposition Electoral Systems." *Electoral Studies* 20(1): 63–86.
- Holland, Paul. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81(396): 945-970.
- Keele, Luke J. "Rbounds: An R Package for Sensitivity Analysis with Matched Data."
- Keele, Luke J. 2009. "An Overview of Rbounds: An R Package for Rosenbaum Bounds Sensitivity Analysis with Matched Data." Retrieved from <http://www.polisci.ohio-state.edu/faculty/lkeele/rbounds%20vignette.pdf>
- King, Gary and Langche Zeng. 2006. "The Dangers of Extreme Counterfactuals", *Political Analysis*, 14(2), 131-159.
- King, Gary and Langche Zeng. "When Can History be Our Guide? The Pitfalls of Counterfactual Inference", *International Studies Quarterly*, 51(1), 183--210.
- Lijphart, Arend. 1986. "Proportionality by Non-NR Methods: Ethnic Representation in Belgium, Cyprus, Lebanon, New Zealand, West Germany and Zimbabwe". In Bernard Grofman and Arend Lijphart ed. *Electoral Laws and Their Political Consequences*. New York: Agathon Press.
- Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press
- Maeda, Ko. 2008. "Re-examining the Contamination effect of Japan's Mixed Electoral System Using the Treatment-Effects Model." *Electoral Studies* 27(4): 723–731

- Mizusaki, Tokifumi and Hiroki Mori. 1998. "Tokuhyou deeta kara mita heiritsusei no mekanizumu." *Senkyo Kenkyu [Electoral Studies]* 13: 50–59.
- Reed, Steven R. 2001. "Duverger's Law Is Working in Italy." *Comparative Political Studies*. 34(3): 312-27.
- Reed, Steven R. 2003. "Heiritsusei ni okeru shousenkyoku kouhosha no hireidaihyou tokuhyouritsu eno eikyou" [The Effect of the PR Tier on SMD Nomination Strategies in Mixed Member Systems: Some Evidence from Japan]. *Senkyo Kenkyu [Electoral Studies]* 18: 5–11.
- Reed, Steven R. 2007. "Duverger's Law is Working in Japan." *Senkyo Kenkyu [Electoral Studies]*. 22: 96-106.
- Rosenbaum, Paul R. 2002. *Observational Studies*. Springer Series in Statistics, New York, Springer. Second edition.
- Rosenbaum, Paul R. 2005. "Sensitivity Analysis in Observational Studies" In Brian S. Everitt and David C. Howell ed. *Encyclopedia of Statistics in Behavioral Science*, New York: John Wiley and Sons, 1809-1814.
- Rosenbaum, Paul R. and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1): 41–55.
- Roy, Andrew D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3: 135-146.
- Rubin, Donald B. 1980. "Bias Reduction Using Mahalanobis-Metric Matching." *Biometrics* 36(2): 293–298.
- Rubin, Donald B. 1986. "Statistics and Causal Inference: Comment: Which Ifs Have Causal Answers." *Journal of the American Statistical Association* 81(396): 961-962.
- Shugart, Matthew S., and Martin P. Wattenberg ed. 2001. *Mixed-member Electoral Systems: the Best of Both Worlds?* Oxford, UK: Oxford University Press.
- Sekhon, Jasjeet. in press. "Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R." *Journal of Statistical Software*. Retrieved from <http://sekhon.berkeley.edu/papers/MatchingJSS.pdf>
- Taagepera, Rein and Matthew S. Shugart. 1989, *Seats and Votes: The Effects and Determinants of Electoral Systems*. New Haven: Yale University Press.

Wildavsky Aaron B. 1959. "A Methodological Critique of Duverger's Political Parties"  
*The Journal of Politics*, 21(2): 303-318

Yitzhaki, Shlomo. 1996. "On Using Linear Regressions in Welfare Economics." *Journal of Business & Economic Statistics* 14(4): 478-486

## TABLES

Table 2.1: Types of Treatment

<i>Candidate</i>	<i>Nominate (treated) in t1</i>	<i>Not Nominate (controlled) in t1</i>
<i>Yes (treated) in t0</i>	<i>Keep (treated -&gt; treated)</i>	<i>Exit (treated -&gt; controlled)</i>
<i># of dist of the SDPJ 2003/ 2005/ The JCP 2009</i>	<b>29 / 26/ 150</b>	<b>25 / 35/ 122</b>
<i>No (controlled)in t0</i>	<i>Participate (controlled -&gt; treated)</i>	<i>Keep Out (controlled -&gt; controlled)</i>
<i># of dist of the SDPJ 2003/ 2005/ the JCP 2009</i>	<b>19 / 12 / 1</b>	<b>169 / 224 / 24</b>

Table 2.2: Estimates of Contamination Effect

	<i>treatment effect (%)</i> <i>(std.err.)</i>	<i># of treated/</i> <i>observation</i>	<i>Sensitivity</i> <i>Analysis</i> <sup>*1</sup>
<i>The SDPJ in 2003</i>			
<i>GenMatch (ATT)</i>			
<i>Nominate/Not Nominate (Pooled)</i>	2.11(.39) ***	48/242	5.0
<i>Keep/ Exit (w/ candidate in 2000)</i>	2.12(.63) ***	29/54	4.0
<i>Participate/ Keep Out (w/o cand in 2000)</i>	2.20(.42) ***	19/188	5.0
<i>GenMatch (ATE)</i>			
<i>Nominate/Not Nominate (Pooled) <sup>*2</sup></i>	1.95(.29) ***	48/242	18.0
<i>OLS</i>	1.70(.20) ***	48/242	
<i>Treatment-Effects Model</i>	1.60(.78) *	48/242	
<i>The SDPJ in 2005</i>			
<i>GenMatch (ATT)</i>			
<i>Nominate/Not Nominate (Pooled)</i>	2.65(.67) ***	38/297	6.0
<i>Keep/ Exit (w/ candidate in 2003)</i>	3.65(.99) ***	26/61	5.5
<i>Participate/ Keep Out (w/o cand in 2003)</i>	.86(.75)	12/236	n.a.
<i>GenMatch (ATE)</i>			
<i>Nominate/Not Nominate (Pooled) <sup>*2</sup></i>	1.31(.43) ***	38/297	7.0
<i>OLS</i>	1.88(.23) ***	38/297	
<i>Treatment-Effects Model</i>	1.07(1.17)	38/297	
<i>The JCP in 2009</i>			
<i>GenMatch (ATT)</i>			
<i>Keep / Exit (w/ cand in 2005)</i>	.88(.37) *	122/272	4.5

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

\*1 Minimum value of  $\Gamma$  required to achieve  $p$ -value over .05 (by .5)

\*2 Matching approach fails to achieve good covariate balance between the treated and the controlled

## APPENDIX: GENERALIZED MAHALANOBIS DISTANCE

The distance is measured with the following distance matrix:  $X$  is observed covariates,  $S$  is the sample covariate matrix of  $X$ , and  $W$  is a  $k$  by  $k$  positive definite weight matrix and  $S^{-1/2}$  is the Choleskey decomposition of  $S$  which is the variance- covariance matrix of  $X$ .

$$d(X_i, X_j) = \{(X_i, X_j)' (S^{-1/2})' W S^{-1/2} (X_i, X_j)'\}^{-1/2}$$

The generic matching is different from traditional Mahalanobis distance approach in the weight matrix  $W$ . Genetic matching approach sets  $W$  to improve the balance in the covariates between the treated and untreated units. Because the approach uses multi- dimensional weighted distance of the covariates to find the match, the model depends less on the parametric assumption among the covariates, treatment and the outcome than another widely used matching approach, the propensity score matching with the estimated propensity score for the covariate balancing (Rosenbaum and Rubin, 1983). See Sekhon (Forthcoming) for more detail.

## Chapter 3

### **Moving the Mountain: Strategic Small-Party Nominations to Shift Major-Party Policy Position under Plurality Electoral System**

#### **INTRODUCTION**

This chapter demonstrates a practical purpose behind small party nominations under plurality electoral rule with two major centrist parties. More specifically, I argue that the small party uses the nomination of candidates to blackmail the major centrist party to adopt policies more desirable to the small party, despite the risk of splitting votes and benefiting an ideological enemy. The small party may not win office, seats, or even votes, which are generally assumed as purposes of political parties (Müller and Strøm 1999). However, the party can achieve more desirable policy outcomes through a position shift of the major centrist party. The major party acts strategically to decrease the risk of nomination and potential electoral damage incurred by the nomination. This argument stands in contrast to previous studies regarding the small parties' nomination that ascribe a purpose of nomination to ideological and/or long-term goals (Cox 1997).

In the first part of the chapter, I present a model based on the *chain-store paradox* (Selten 1978; Kreps and Wilson 1982), in which a small radical party is able to exert policy influence in repeated negotiations with a major centrist party located on the same side of the ideological median. The model indicates that if the major party believes that there is even a small chance that the small party is *ideological* -- it prefers the defeat of the major party to victory, if the major party chooses too centrist position -- a less *ideological*, more *policy oriented* small party is able to persuade the major party to choose a less

centrist position by using the nomination threat. In the second part of the chapter, I show empirical evidence that a small leftist party and the major center-left party in Japan choose their strategies in accordance with the model.

For decades, political scientists have studied how the heterogeneous interests of electorates are represented in policy. Many scholars have focused on the electoral system, because their rules have a large impact on the process of representation, via the shape of the party system, the form of the government (single party or coalition), and the choice that the electorates face at elections (Gallagher and Mitchell 2006).

Traditionally, electoral systems are separated into two broad categories: single-member plurality and proportional representation (PR) systems (Grofman and Lijphart 1986). In his seminal work, Downs establishes the starting point for models of party competition and subsequent representation under different electoral systems. The Downsian model predicts that with two-party competition under the plurality electoral system, the issue positions of the two parties converge to that of the median voter. On the other hand, under a PR system with multiple parties, parties are expected to have considerably different policy positions (Downs 1957). In other words, the Downsian model implies that it could be more difficult for radical groups to have their voice heard in the legislative process under plurality electoral systems than under PR.

It is true that this outcome depends on many theoretical assumptions of the basic Downsian spatial model (Grofman 2004). If there is a potential entry of third parties into the contest, even under the plurality electoral system the two established parties could select spatially separated positions - more attentive to the preferences of the radical constituencies - to deprive small parties of a chance of victory (Palfrey 1984).

However, other assumptions help support the original outcome even if the third parties might enter the contests. If the electorate votes strategically, not necessarily sincerely, they might be less inclined to vote for new party candidates, even if they have more agreeable policy position. They use the endorsement of the major established party as a focal point for their strategic coordination, and desert new entrants and/or small party candidates because they are expected to be less viable (Cox 1997). Moreover, the



established parties often have other non-policy advantages, such as qualified and experienced personnel, developed party organization, and access to government resources. Potential entrants can be also discouraged by disproportional seat to vote-share provision under plurality electoral rules. Hence, once two major parties establish their dominant status, and seek to occupy the centrist position as the Downsian model implies, small radical parties or groups have hard time influencing the policy.

Anticipating these disadvantages for the third parties, political parties - especially smaller parties -- have incentives to form pre-electoral alliances with some policy accords if there are more than two parties competing under plurality electoral rule (Golder 2006). The subsequent bipolar alliance sometimes results in the formal unification of the parties, although there is a cross-national variation in how the district-level bipolarity develops to the national bipartisanship (Chhibber and Kollman 2004).

Nevertheless, small parties often nominate their party candidates to plurality contests without a pre-electoral alliance. These nominations seem not to have any practical purpose. First, these small parties have little chance of getting their candidate elected most of the time. Moreover, the nominations could be even harmful to their own interest if they care about the policy outcome chosen by the campaign winner. Their nominations might help the candidates of an ideologically distant major party win by seizing some votes from the candidates of the ideologically-closer major party. Therefore, previous studies discuss that the purpose of their nominations is ideological and/or long-term oriented (Cox 1997), such as representation of ethnic minority or maintenance of a distinct political ideology, hoping that their opinion will prevail in the long run.

This chapter, on the other hand, demonstrates an alternative, *practical* logic behind small parties' independent campaign<sup>13</sup>. In this chapter, *practical* means that the

---

<sup>13</sup> The other potential purpose behind nominations is the contamination effect. Parties may coordinate their campaign strategy among the different electoral tiers of the same election, different levels of the elections and the elections of the different legislative chambers because of their spillover effects: the *contamination* effect. The effect has been discussed mainly in the context of small parties campaigning independently on the

actor chooses the instrumentally rational strategy to achieve their goal as defined in the model, such as a more preferable policy outcome. On the other hand, I use the term *ideological* as the antonym of *practical*: *ideological* actors might not select the instrumentally rational strategy to accomplish the goal, such as punishing ideological closer party, though it could benefit an ideological enemy.

Small parties nominate candidates to plurality contests if the major parties on their side of the ideological median choose positions are deemed “too centrist”. This strategy seems to pursue very *ideological*, not *practical* goals because the nomination has little chance of yielding a seat, and also could benefit their ideological enemies in the contest. However, I argue that the nomination serves a very *practical* goal – a more preferable policy outcome, which is achieved by the policy concession of the major centrist parties towards the ideal point of the small parties. To avoid the participation of the small party in the contest, the major party chooses a less centrist position and consequently the small party achieves a more preferable policy outcome, even if the party does not actually gain seats in the legislature<sup>14</sup>.

This chapter is organized as follows. In the first part of the chapter, I discuss a model that describes the repeated games between a center-left major party and a small leftist party over nominations. Next, I test hypotheses inferred from the model with empirical data from Japan. The final section discusses the implications of the theoretical and empirical results.

## ASSUMPTIONS, DEFINITIONS AND GAME SEQUENCE

This section describes the model of the strategic interactions between a small radical party and a major centrist party located on the same side of the ideological median. To

---

plurality tier of the mixed-member electoral systems by aiming to mobilize more support in the proportional tier (Cox and Schoppa 2002; Ferrara et al. 2005; Herron and Nishikawa 2001; Reed 2003; Maeda 2008). Other studies also suggest the effect is working between national level elections in plurality electoral rule and sub-national level elections with more proportional electoral rules (Horiuchi and Natori 2007)

<sup>14</sup> Cox briefly mentions the possibility of this *blackmailing* strategy of the small parties (pp162, Cox 1997), although he does not discuss the strategy in detail.

achieve a policy concession from the major centrist party, the small party uses the threat of the nomination to blackmail the major party. The model of this chapter, which applies *the chain-store paradox* (Selten 1978; Kreps and Wilson 1982) shows that in repeated interactions, the small party can convince the major party that it is better off by choosing less centrist position to avoid the nomination, if there is even a small probability that the small party is *ideological*.

In the *chain-store paradox* game, a monopolist faces a succession of potential entrants over demarcated local markets, which can be modeled as finite repeated games between the actors. Game theoretical reasoning, which Selten terms as *the induction theory*, implies that the monopolist does not fight a price war with the entrants to maximize its utility. Nevertheless, Kreps and Wilson show that if the potential entrants are uncertain about the monopolist's payoffs, the monopolist chooses to fight a price war in the early stages of negotiation to establish a reputation as *vengeful*, which Selten terms as *the deterrence theory*. Thus, the potential entrants in later stages avoid challenging the monopolist even if the monopolist's payoff is worse in the market due to the price war. The authors also indicate the outcome does not change if a single rival negotiates over repeated opportunities to enter.

Tsebelis refers to similar strategic interactions in the British Labour Party between the MPs and local activists in the 1970s (Tsebelis 1990). The local Labour activists, who tend to hold more leftist views than regular voters or the MPs, could deter the re-nomination of incumbent MPs if they judge the MPs as not sufficiently leftist, even though the replacement of the candidates hurts the Labour's chance to defeat the Tory party. Tsebelis argues that, though it seems irrational for the activists to replace the moderate incumbent MPs, the equilibrium could support replacement under uncertainty and iteration of the game.

Despite the implications of the game on various topics, to the best of my knowledge there is no subsequent work on the games between radicals and moderates regarding the position of candidates that compete under plurality electoral rule. Empirical studies that test the model are similarly lacking. The model in this chapter

refers to Tsebelis's model, though the details are different based on the different strategic environments between the parties.

### Actors and Game Sequence

This chapter starts with a simple extensive game with two players: a major center-left party, *the Democrats (D)*, and a small radical-left party, *the Socialists (S)*, and two districts ( $d = 1, 2$ ) that elect a representative with a first-past-the-post rule. For both parties, the national headquarters, not the local party apparatuses, have authority for their nomination<sup>15</sup>.

There is another major party, the center-right party, the *Conservatives*. The two major parties, the Democrats and the Conservatives always participate in both districts. However, prior to the interactions between the Democrats and the Socialists, the Conservatives unilaterally choose the policy position on the right side of the ideological median. Hence, the Conservatives are excluded from the game because their strategy is exogenously given.

The Democrats and the Socialists sequentially choose their strategies in  $d = \{1, 2\}$ . For each district, the Democrats first decide whether the party chooses the centrist (C) or the leftist (L) position. The Socialists on the other hand choose between participate/nominate (N) or acquiesce/do not nominate (A) after the party observes the choice of the Democrats. The Socialists do not have to pay any cost for the nomination, but their candidates do not have a chance to get elected under the plurality electoral rule. Moreover, the party might fail to recruit candidates with some probability  $\varepsilon$ , such that  $\varepsilon < 1/2$ , because of their limited organizational resources. Figure 3.1 shows the sequence of the model.

The Socialists can be either of two types, *Ideological (I)* or *Practical (P)*. The *Ideological* type hates the centrist Democrats as traitors to their leftist ideal, and always

---

<sup>15</sup> It is possible to relax the assumption for the Democrats. The outcome does not change even if the local party organizations of the Democrats decide their strategy individually across the districts, as the original *chain-store paradox* model assumes for the entrants.

attempts to nominate against them even if it could benefit their ideological enemy, the Conservatives. The *Practical* is more modest and policy oriented: they understand that they cannot attain seats even if they nominate, but may run a candidate against the centrist Democrats by aiming that the nomination produces more preferable policy outcomes to the party. The Democrats initially have a belief that the Socialists are *Ideological* with the probability  $\delta_1$ , such that  $\delta_1 < 1/2$ .

The *Ideological* Socialists correspond to the *vengeful* monopolist in Kreps and Wilson's *chain-store paradox*. The difference between my model and their model is that they assume the *vengeful* monopolist always fights the entrant, but I assume that the *Ideological* Socialists always attempt to nominate against the centrist Democrats, but they may fail to recruit candidates with the probability  $\varepsilon$ . Similarly, the *Practical* Democrats also might fail to run the candidates with the same probability  $\varepsilon$ , even if they attempt to. From the Democrats' point of view, the term  $\varepsilon$  gives some uncertainty between the Socialist will and behavior.

### Preference

This chapter assumes that the Democrats are office-oriented i.e. party gains a payoff 1 if it wins, 0 if it loses; the party is indifferent to the party position as long as it wins. On the other hand, the Socialists are concerned with the type of Democrat who wins the district; the Socialists attain payoff  $L$  if the leftist Democrats win,  $C$  if the centrist Democrats win, and 0 if the Democrats lose to the Conservatives.

The chance of the Democrats' victory depends on the policy position of the Democrats and the Socialists' strategy. As the Downsian model infers, the centrist Democrats have a better chance of winning the election, owing to their proximity to the electoral median, if the other conditions are held constant. However, the Democrats have a smaller chance to win if the Socialists participate in the campaign. The likelihoods of the Democrats victory are  $b_N$ ,  $b_L$  or  $b_A$ , where  $b_N = P(\text{Democrats win} \mid \text{Democrats run centrist, Socialists nominate})$ ,  $b_A = P(\text{Democrats win} \mid \text{Democrats run centrist, Socialists acquiesce})$ , and  $b_L = P(\text{Democrats win} \mid \text{Democrats run leftist (\& Socialists acquiesce)})$ . This chapter assumes that

$0 < b_N < b_L < b_A < 1$  and  $(b_A - b_L)/(b_A - b_N) < 1/2$ , and the respective likelihoods are the same between  $d=1$  and 2.

Because the Democrats attain a payoff 1 if the party wins, their expected payoffs for respective outcomes are the same with their likelihoods of victory,  $b_N$ ,  $b_L$  or  $b_A$ . On the other hand, for the Socialists, the expected payoffs are the products of the payoffs of each Democrats type and the likelihood of her victory: the Socialists attain  $b_LL$  if the Democrats select the leftist,  $b_NC$  if the party chooses to fight and gain  $b_AC$  if acquiescing, given the Democrats run the centrist. Both types of the Socialists prefer most that the leftist Democrats win. The *Practical* Socialists prefer a centrist win to a Democrat loss;  $L > C > 0$ , though the *Ideological* type prefers the centrist loses;  $L > 0 > C$ . The *Practical* type also prefers the Democrats choose a leftist more than the other outcomes:  $0 < b_NC < b_AC < b_LL$ , and the difference in the payoffs between  $b_LL$  and  $b_AC$  is larger than those between  $b_AC$  and  $b_NC$  (i.e.  $b_LL - b_AC > b_AC - b_NC$ , or  $b_LL + b_NC > 2b_AC$ ).

### Strategies and Beliefs

The Democrats move first and select the position at  $d=1$ , the centrist or the leftist, based on the expected payoffs of each outcome. The Democrats expect that the Socialists choose  $N$  with the probability  $p_1 = \delta_1(1-\varepsilon) + (1-\delta_1)q(1-\varepsilon) = \{\delta_1 + (1-\delta_1)q\}(1-\varepsilon)$  if they nominate the centrist, given their belief  $\delta_1$ , the probability of the *Practical* Socialists to nominate  $q$  (explained more detail below), and the likelihood of the recruiting failure  $\varepsilon$ . Therefore, the Democrats choose the leftist if  $b_L > p_1 b_N + (1-p_1)b_A$ , where  $p_1 > (b_A - b_L)/(b_A - b_N)$ . They run the centrist if  $p_1 < (b_A - b_L)/(b_A - b_N)$ , and are indifferent if  $p_1 = (b_A - b_L)/(b_A - b_N)$ . Let  $(b_A - b_L)/(b_A - b_N) = b^*$ , such that  $0 < b^* < 1/2$ .

If the Democrats select the centrist position, the Socialists choose between  $N$  and  $A$ . The Democrats update the belief  $\delta_1$  to either  $\delta_N$  or  $\delta_A$ , based on the Socialist reaction to the centrist Democrats, according Bayes rule. If the Democrats choose the leftist position at  $d=1$ , the initial belief  $\delta_1$  does not change, because the party attains no additional information. If the Socialists choose  $N$  at  $d=1$ , the Democrats' belief at  $d=2$  is updated to  $\delta_N$  as follows:

$$\begin{aligned}
\delta_N &= P(\text{the Socialists are Ideological} \mid \text{Nominate in } d=1) \\
&= \delta_I(1 - \varepsilon) / [\{\delta_I + (1-\delta_I)q\}(1-\varepsilon)] \\
&= \delta_I / \{\delta_I + (1-\delta_I)q\}
\end{aligned}$$

Similarly, if the Socialists choose  $A$ , the belief is updated to  $\delta_A$  as follows:

$$\begin{aligned}
\delta_A &= P(\text{the Socialists are Ideological} \mid \text{Acquiesce in } d=1) \\
&= \delta_I \varepsilon / [\delta_I \varepsilon + (1-\delta_I)q(1-\varepsilon) + (1-\delta_I)(1-q)] \\
&= \delta_I / [\delta_I + (1-\delta_I)q(1-\varepsilon)/\varepsilon + (1-\delta_I)(1-q)/\varepsilon] < \delta_I
\end{aligned}$$

In  $d=2$ , the Democrats select the position based on the updated belief of the Socialists  $\delta$ , such that  $\delta_N$ ,  $\delta_A$  or  $\delta_I$ . The Socialists who are *Practical* always choose  $A$  in  $d=2$  if the Democrats choose the centrist, because  $A$  is always a better strategy than  $N$  if there is no following bargain. Given the Socialists' strategy, at  $d=2$  the Democrats expect that the Socialists nominate with the probability  $p_2 = \delta(1 - \varepsilon)$ : only the *Ideological* type attempts to nominate. Therefore, the Democrats choose the centrist if  $p_2 < b^*$ , or  $\delta < b^*/(1 - \varepsilon)$ .

The *Ideological* Socialists always (attempt to) nominate if the Democrats run the centrist in both districts. The *Practical* type always chooses  $A$  in  $d=2$ , but at  $d=1$  it considers the impact of the decision to the Democrats belief  $\delta$ , which in turn can change the Democrats' choice in  $d=2$ .

If the Democrats choose the centrist in  $d=1$ , which implies that  $p_1 < b^*$ , and the Socialists acquiesce in response, the Democrats surely nominate the centrist again at  $d=2$ , because  $p_2 = \delta_A(1-\varepsilon) < \delta_I(1-\varepsilon) < \{\delta_I + (1-\delta_I)q\}(1-\varepsilon) = p_1 < b^*$ . Hence, if the *Practical* Socialists acquiesce (or fail to recruit the candidates) in  $d=1$ , they receive  $2b_{AC}$  in total. On the other hand, if they nominate in  $d=1$ , their expected payoff is  $b_{NC} + b_{LL}$  if  $p_2 = \delta_N(1-\varepsilon) > b^*$ , but  $b_{NC} + b_{AC}$  if  $\delta_N(1-\varepsilon) < b^*$ . The *Practical* Socialists choose their strategy  $q$  in  $d=1$  given these payoffs. The payoff is  $q(b_{NC} + b_{LL}) + (1-q)2b_{AC}$  if  $\delta_N(1-\varepsilon) > b^*$  and  $q(b_{NC} + b_{AC}) + (1-q)2b_{AC}$  if  $\delta_N(1-\varepsilon) < b^*$  according the party's choice of  $q$ .

### Equilibrium

Equilibrium of the model depends on  $q$ : how likely it is that the *Practical* Socialists nominate if the Democrats run centrist candidates. All the other related factors,  $\delta_1$ ,  $b^*$  and  $\varepsilon$  are defined exogenously.

$q = 1$  is not the equilibrium: If  $q = 1$ ,  $\delta_N = \delta_1 / \{\delta_1 + (1-\delta_1)\} = \delta_1$ . Hence, if the Democrats run the centrist in  $d=1$  (i.e.  $p_1 < b^*$ ), the party also nominates the centrist in  $d=2$  even if the Socialists select  $N$  in  $d=1$ , because  $p_2 = \delta_N(1-\varepsilon) = \delta_1(1-\varepsilon) < p_1 < b^*$ . In other words, if the Democrats know that not only the *Ideological* but the *Practical* Socialists always attempt to choose  $N$ , the nomination is not persuasive evidence that the Socialists are *Ideological*. Therefore, both the Socialists and the Democrats *know* that the *Practical* Socialists should not choose  $q = 1$ . On the other hand,  $q = 0$  is also not the equilibrium: If  $q = 0$ ,  $\delta_N = \delta_1 / \{\delta_1 + (1-\delta_1)0\} = 1$ . The *Practical* Socialists have some incentive to run (i.e.  $q > 0$ ), because the Democrats surely nominate the leftist in  $d=2$  if the Socialists nominate at  $d=1$ , given  $p_2 = \delta_N(1-\varepsilon) = (1-\varepsilon) > 1/2 > b^*$ . Consequently, not only the Socialists but also the Democrats understand that the *Practical* Socialists do not set  $q$  at 0.

The *Practical* Socialists' payoff monotonously increases with the value of  $q$  if  $\delta_N(1-\varepsilon) > b^*$  because  $(b_{NC} + b_{LL}) > 2b_{AC}$ , but monotonously decreases if  $\delta_N(1-\varepsilon) < b^*$  owing to  $b_A > b_N$ . Moreover,  $q(b_{NC} + b_{LL}) + (1-q)2b_{AC} > q(b_{NC} + b_{AC}) + (1-q)2b_{AC}$  for all  $q$  such that  $0 < q < 1$  because  $b_{LL} > b_{AC}$ . Consequently, in equilibrium the *Practical* Socialists randomize their response  $q$ , at the point satisfying  $\delta_N = \delta_1 / \{\delta_1 + (1-\delta_1)q\} = b^* / (1 - \varepsilon)$ . By solving the equation,  $q = [\delta_1 / (1-\delta_1)] * [1 - (b^* + \varepsilon) / b^*]$ .  $q$  has a unique solution between 0 and 1, given the conditions that  $\delta_1 < 1/2$ ,  $b^* < 1/2$  and  $\varepsilon < 1/2$ . At this point, the *Practical* Socialists can choose  $N$  as frequently as possible to gain benefits from the shift of the Democrats' strategy. The Democrats understand that not only the *Ideological* but *Practical* Socialists attempt to nominate with the probability  $q = [\delta_1 / (1-\delta_1)] * [1 - (b^* + \varepsilon) / b^*]$ , but still better off by choosing the leftist in  $d=2$ .

Therefore, the equilibrium is given as follows: If  $b^* < \delta_1(1-\varepsilon)$ , the Democrats choose the leftist candidates in both  $d = 1$  and 2. If  $\delta_1(1-\varepsilon) < b^*$ , the Democrats run the centrist in  $d=1$ . The *Ideological* Socialists always attempt to nominate if the Democrats select the centrist, and the *Practical* Socialists attempt to nominate with the probability  $q$ ,



which satisfies that  $q = [\delta_1/(1-\delta_1)]*[(1 - (b^* + \epsilon))/b^*]$ . If the Socialists nominate in  $d=1$ , the Democrats choose the leftist at  $d=2$ . If the Socialists do not nominate, the Democrats run the centrist at  $d=2$ . The *Ideological* Socialists nominate against, and the *Practical* type acquiesces in response at  $d=2$ .

### Implication & Extension

This model demonstrates that if the Democrats believe that there is some small chance that the Socialists are *Ideological*, the *Practical* Socialists may run candidates against the centrist Democrats to attain more preferable policy outcome in the subsequent negotiations. Although previous studies argue that the nomination of the small parties in plurality contest is ideological or long-term oriented (Cox 1997), the model shows that small parties have incentive to nominate even if they have *practical* preference, i.e. to achieve a policy concession from the major party.

As another implication, this model also suggests that the *Practical* Socialists are no more likely to nominate even if the electoral result is foreseeable, counter to the argument by Tsebelis (1990). Tsebelis argues that the local constituency is more likely to replace the moderate MP if the district is uncompetitive: either safe Labour or safe Tory, because the replacement does not change the electoral outcome (pp130- Tsebelis 1990). If his argument is applied to the model in this chapter, the *Practical* Socialists should be more likely to choose  $N$  at  $d=1$  if the electoral outcome is predictable:  $b_A - b_N$  is smaller, because the Socialists can send the signal to the Democrats without hurting the Democrats chance of victory.

However, the model suggests the other ways around. If the Socialists nomination is less likely to change the Democrats' chance of victory ( $b_A - b_N$  becomes smaller),  $b^* (= (b_A - b_L)/(b_A - b_N))$  goes up. As the term  $b^*$  is larger, then  $q$  in the equilibrium becomes *smaller* because  $q = [\delta_1/(1-\delta_1)]*[(1 - (b^* + \epsilon))/b^*]$ :  $q$  is monotonously decreasing function with  $b^*$  within its range. Hence, the model implies that the *Practical* Socialists are less likely to nominate if their nomination does not change the electoral result.

This model can be easily extended to the environment with more than two districts. If the parties play the game with more than two districts, the *Practical* Socialists would evaluate the consequences of the decision not only based on the Democrats' belief in the next district, but also at all of the districts to be played between them in the future. If the game is iterated several times, the *Practical* Socialists would have more incentive to fight, especially in the early phase of the iteration, because that decision affects the negotiation in all of the subsequent districts.

The other possible extension is the incumbent candidates. The model assumes that the Democrats can choose either the leftist or the centrist position without cost. However, if the party has incumbent candidates, it might be costly if trying to move from the original position. If a candidate changes her position, her core constituency would take it as a betrayal to their cause. Similarly, if the party replaces a candidate for the policy reason, core supporters of the former candidate could defect.

If *Nature* chooses the centrist as the Democrats' incumbent,  $b^* = (b_A - b_L) / (b_A - b_N)$  will be larger because the Democrats have less chance to get their candidate elected if the party replaces the candidates from the leftist to the centrist:  $b_L$  is smaller. Therefore, the Democrats are more likely to choose the centrist given that  $p_1$  and  $p_2$  are held constant. On the other hand, if *Nature* chooses the leftist, it likely leads to smaller  $b^*$  because  $b_A$  is smaller: the Democrats choose the leftist given the other conditions are constant. In other words, the Democrats are less sensitive to the perceived risk of the Socialists' nomination in the districts with their incumbent candidates than those in open districts.

## EMPIRICAL ANALYSES

### Empirical Background of the Case and the Data

In this section, I empirically test the theoretical model of the strategic interactions between the major center-left party and the small leftist party discussed in the previous section. Because of the limitation on the observable data, in this section I focus on examining the following two decisions: 1) whether the leftists (i.e. the Socialists) choose

to nominate to send the signal for their type if the center-left party (i.e. the Democrats) run the centrist, and 2) whether the center-leftist party preemptively choose the type of the candidates according to the support level of the leftist party, which influences their risk of nomination and the level of the damage incurred by the nomination.

For the tests, I use electoral data and candidate positions from the Japanese House of Representatives (HR) election in 2003 and 2005. These elections provide an ideal case for testing the model discussed in the previous section. I will briefly describe the background of the case, especially why the campaigns in the plurality electoral tier could be simplified into the sequential strategic interactions between the Democratic Party of Japan (DPJ) and the Social Democratic Party of Japan (SDPJ) -- the *Democrats* and the *Socialists* -- two leftist parties in Japan.

The Japanese electoral system has used the first-past-the-post rule for one of the two electoral tiers of the mixed-member majoritarian (MMM) electoral system since 1994. A Japanese elector has two votes, one is for a candidate in her single member district that elects a representative with the first-past-the-post rule (plurality tier), and the other is for a party list, which provides seats according to proportional rule (the proportional representation (PR) tier). In the MMM system, the votes cast to one electoral tier do not have influence on the outcome of the other tier<sup>16</sup>. The Japanese voters experienced five elections under the current system, in 1996, 2000, 2003, 2005 and 2009. Despite several opposing predictions, the plurality tier of MMM system appears to give the same Duvergerian pressure as the pure plurality system: decreasing the number of candidates competing at each district (Reed 2001; 2007). Therefore, this chapter uses the plurality tier of the Japanese HR elections as the case to test the model of strategic interaction over nomination.

---

<sup>16</sup> There is another type of the mixed-member system used in Germany and New Zealand. The system – the mixed-member proportional (MMP) system – decides the allocation of the seats based on the votes cast to the proportional tier. The MMP system uses the plurality tier to decide which candidates gain priority among the party list, given the number of seats allocated to the party. Therefore, the system functions as a proportional electoral system, except that the system gives priority to the candidates popular in their districts.

The Japanese party system has been fluid since the split of the Liberal Democratic Party (LDP) in 1993. The present party system - two major and a few smaller parties -- was established in 2003, when the Liberal Party merged with the DPJ. In other words, since the 2003 election, Japanese voters are faced with two major political camps competing for power. The two major parties of Japan - the LDP on the center-right and the DPJ on the center-leftist - have quite different histories. The LDP has been predominant in Japanese politics since 1955. The party held power continuously for more than a half century since its formation until 2009, except for a short break in 1993. The LDP corresponds to the *Conservatives* in the model.

The DPJ, by contrast, was formed in 1996 by the center-to-leftist politicians who were members of the Japanese Socialist Party (JSP) and other small center-left parties. The DPJ, however, obtained several conservative politicians after its formation. First, many of the former New Frontier Party (NFP) members, the largest opposition, joined the DPJ after the party's breakup in 1997. The Liberal Party, a small conservative party led by the ex-LDP cadre Ichiro Ozawa, merged with the DPJ in 2003 before the HR election. The new members from the NFP and the Liberal Party have quite different political backgrounds from those of the original DPJ members. Hence, the DPJ consists of politicians with heterogeneous political views, especially in the foreign/defense issues that divided Japanese politics during the Cold War era.

The DPJ usually re-nominates incumbent candidates in subsequent elections except in case of retirement or extraordinary situations such as corruption scandal. To the best of my knowledge, the party has never chosen not to renominate an incumbent candidate because of her policy position. If the party selects a new candidate to a district, it does so with the following process. First, the prefectural chapter nominates a potential candidate to the national headquarters under consultation with the district level chapter and the prefectural party executive board<sup>17</sup>. At the national level, the national party

---

<sup>17</sup> For example, refer to the Article 17, the Platform of the DPJ Kanagawa Prefecture Chapters Confederation, which is available from <http://www.dpjr.org/organization/rules/index.html>

executive board considers the selection, and finally the standing officer council approves the selection<sup>18</sup>. In other words, although the national headquarters has a final veto over the nomination, the local party organization has considerable voice in the process<sup>19</sup>.

For the elections used for the analysis – the HR elections in 2003 and 2005 -- most of the DPJ incumbent candidates had started campaigning in their district before the current party system was established in 2003. In other words, the DPJ candidates chose their position or were selected owing to their policy position before the two major parties and some smaller parties began competing in the current environment. Senior DPJ politicians began their career as the members of the HR under the previous SNTV/MMD electoral system. Other junior candidates began running as candidates of the NFP that tended to be more rightist than the current DPJ, or as DPJ candidates who were more left-leaning because there were two other large parties, the LDP and the NFP on the rightist side of the policy space. Because of the different party configuration before 2003, the DPJ incumbent positions are not expected to be well-fitted to the current political environment of each district, especially with respect to the SDPJ's threat.

In 2003 and 2005, the SDPJ nominated a considerable number of candidates in plurality district without forming formal national-level electoral alliances. The SDPJ consists of leftist members of the former JSP, who did not want or were refused to join the DPJ in 1996. The SDPJ is usually perceived as being located to the left of the DPJ, and the right of the Japanese Communist Party (JCP). During 2003 and 2005, the SDPJ had an informal cooperative relationship with the DPJ in legislative process and in campaigns at some districts. However, the party ran their candidates in other districts where the DPJ also nominated; hence the SDPJ's candidates competed with the DPJ in those districts<sup>20</sup>.

---

<sup>18</sup> *The DPJ Platform*, the Article 18. <http://www.dpj.or.jp/governance/policy/>

<sup>19</sup> I attained valuable advice about the DPJ's candidate selection process from Prof. Jun Saito, who has served as the DPJ's Representative of the Yamagata fourth district.

<sup>20</sup> In contrast, the SDPJ chose to form an electoral accord in most of the districts with the DPJ for the 2009 election, because the two parties developed cooperation to achieve majority in the House of Councilors (HC) since 2007. Therefore, the 2009 HR election is not appropriate to test the model.

Therefore, the party is ideal to test the main hypothesis of the chapter: a small radical party uses nominations in plurality contests to pressure the major party to choose candidates with more favorable political views.

I briefly describe the other small parties in Japan to explain why the plurality contest in Japan can be simplified as the strategic interaction between the DPJ and the SDPJ. Besides the LDP, the DPJ and the SDPJ, Komei, the JCP, and a few other conservative parties formed by ex-LDP politicians had representatives in the legislature in 2005. However, the process used by these parties does not represent the interaction represented by the model that I described.

First, Komei chose to form a pre-electoral alliance with the major parties. To survive under the electoral system favoring larger parties, the party supports candidates of one of the major parties in the most of the plurality districts, in exchange for the aid in mobilizing PR votes and/or races in a few plurality districts saved for the Komei candidates. Moreover, Komei is not necessarily located on the policy space shared with the other parties. Komei's support base is a Buddhist sect, the *Soka Gakkai*. Though the party is generally assumed to be centrist, it does not have strong appeal to centrist voters beyond their religious base. Hence, the strategy of Komei does not have a strong relationship with the other parties' ideological positioning, and consequently is not appropriate for the model of party competition over the policy space. Next, the JCP – the radical leftist party – chose to campaign alone: the party nominated candidates in most of the districts independently until 2005. Therefore, the JCP's strategy is externally given to the DPJ and there is no strategic interaction between the party and the other parties.

Also, a few small conservative parties were formed before the 2005 election by former members of the LDP, who got ousted from the party due to their opposition to Prime Minister Koizumi's proposal for postal service privatization. However, those parties do not fit well in the "party competition over the issue" dimension. First, the split is based on a single issue that cuts across major parties – postal service privatization. Moreover, the candidates of those parties fought their campaigns based on their personal electoral base established as LDP politicians, not as members of the parties with

established policy platforms. Hence, these small parties are not appropriate to test the model of strategic interactions among political parties over policy space.

I use foreign/defense policy as the issue dimension of Japanese politics for the purposes of this chapter, especially with respect to the division between the DPJ and the SDPJ. Foreign/ defense policy has been the most important political issue dividing the conservative and the leftist camps during the Cold War era (Kabashima and Takenaka 1996) and still has obvious significance among current Japanese representatives (Taniguchi 2006; Laver and Benoit 2006). Moreover, the SDPJ puts considerable importance on the topic, because the party is a remnant of the leftist members of the JSP that sticks to the most leftist policy in the field<sup>21</sup>. Because positions on foreign/defense policy are used in the analysis, in the following I use the terms hawk and dove as the synonyms of the centrist and the leftist of the DPJ candidate. I use data from the University of Tokyo/Asahi Shimbun Survey in 2003 and 2005 to estimate the positions of individual candidates for the analysis<sup>22</sup>. To estimate policy positions in foreign/defense policy, I employ responses to the following seven questions: 1) amendment of the constitution<sup>23</sup>, 2) size of the defense force, 3) the Japan-US alliance, 4) preemptive strike,

---

<sup>21</sup> In June 2010, the SDPJ broke up the coalition government with the DPJ owing to their policy disagreement regarding to a foreign/defense policy, the US military base in Okinawa. This event also suggests the importance of the topic for the SDPJ to non-Japan specialists.

<sup>22</sup> The data are available from <http://www.j.u-tokyo.ac.jp/~masaki/ats/atpsdata.html>. The codebook is included in Kabashima and Yamamoto (2005) for the 2003 survey, and Kabashima and Yamamoto (2008) for 2005. I use *Amelia II* (Honaker, King and Blackwell 2009) to handle the item missing of the responses by the multiple imputation approach. To handle the multiple datasets created by the multiple imputations for the analysis, I use *Zelig* (Imai, King and Lau 2009). For the multiple imputation, I use all responses to political and policy questions within the survey which the answers are measured with ordinal or interval scale. In addition, I create the dummy variables for the DPJ, the SDPJ, the JCP and Komei from the party affiliation of the candidates and add them into the model. For the model simplicity, I use a model that treats all the covariates as being measured with the interval scale. The code used for the analysis is available from the author by request.

<sup>23</sup> In the Japanese political context, *constitutional amendment* often implies the amendment of the Article 9: Renunciation of War, which in effect restricts armament.

5) seeking a permanent seat in the UN Security Council<sup>24</sup>, 6) foreign policy to North Korea, and 7) right of collective defense. The wording of these questionnaires is included in the Appendix.

From the responses to these items, I extract a latent issue dimension of foreign/defense policy with a factor analysis, and then estimate the position of each candidate on the dimension from the responses. Hence, the candidate position is estimated with a continuous scale, though I have discussed in the previous section as if the position is dichotomous -- the centrist and leftist -- for simplicity. The factor explains quite well the variance of the responses in both elections: all of the factor scores are beyond .80, except for the response to the UN Security Council questionnaire in 2005 (the factor score is .53).

### Hypotheses

This section bridges the theoretical and empirical parts of the chapter by drawing hypotheses from the formal model in the previous section, which will be tested via statistical analysis in the next section. First, three hypotheses regarding the SDPJ's strategy is proposed, and then two hypotheses of the DPJ's decision-making are presented.

#### *Hypotheses: the SDPJ's Strategy*

The model in the previous section argues that the issue position of the Democrats is the main factor influencing the decision of the Socialists: Both types of the Socialists do not nominate if the Democrats nominate the leftist, but if the Democrats choose the centrist, the *Ideological* type nominates with the probability of  $1-\varepsilon$ , and the Practical types do with

---

Therefore, responses to the questions are perceived as asking their viewpoints towards their foreign/defense policy.

<sup>24</sup> The Japanese leftists usually oppose Japan seeking permanent seat in the UN Security Council, because they are concerned that Japan would be required to send the army to enforce the UN resolution if Japan is a permanent member of the Council. Therefore, this seemingly neutral policy distinguishes hawks and doves in Japanese politics.



the probability  $q(1-\varepsilon)$ . Therefore, in the empirical analysis first I test that the SDPJ actually run a candidate if the DPJ nominates a centrist to the district. Hence, the hypothesis should be:

*HS1. The SDPJ is more likely to nominate a candidate against the DPJ if the DPJ candidate is a centrist, given the other conditions are held constant.*

On the other hand, Tsebelis implies that the SDPJ will run a candidate if the electoral outcome is predictable, because the SDPJ send a signal about their type without significantly harming the chance of the DPJ's victory (Tsebelis 1990). More precisely, the prediction could be that SDPJ will run the candidates if the DPJ run the centrist *AND* the outcome is predictable. If these are the cases;

*HS2. The SDPJ is more likely to nominate a candidate against the DPJ if the outcome of the LDP-DPJ contests is predictable, given the other conditions are held constant.*

*HS2a. The SDPJ is more likely to nominate a candidate against the DPJ if the DPJ nominate a centrist and the outcome of the LDP-DPJ contests is predictable, given the other conditions are held constant.*

However, the model in the previous section suggests that these hypotheses might not be true. The DPJ might not believe that the nomination as the evidence of the SDPJ being *Ideological* if the electoral outcome is predictable, because the DPJ would expect that not only the *Ideological* but the *Practical* SDPJ would not hesitate to nominate if the result is foreseeable.

#### *Hypotheses: the DPJ's Strategy*

The model in the previous section posits that the Democrats' decision is a function of  $\delta_1$ : their belief on the Socialists' type,  $\varepsilon$ : the likelihood of the Socialists recruitment failure,  $b^*$ : their relative payoffs on respective outcomes, and  $q$ : the Socialists strategy in the equilibrium. In  $d=1$ , the Democrats select the leftist if the risk of the Socialists nomination  $p_1 = \{\delta_1 + (1-\delta_1)q\}(1 - \varepsilon)$  is larger than  $b^*$ , and in  $d=2$  the party runs the leftist if  $p_2 = \delta(1 - \varepsilon)$  is larger than  $b^*$ .

These parameters are not directly observable, especially the subjective beliefs of the DPJ. However, some district attributes expected to be correlated with those parameters are available. This chapter focuses on two factors: the strength of the SDPJ in the districts and the incumbency of the DPJ's candidates.

First, the strength of the SDPJ in the districts is expected to be correlated with  $\delta$ ,  $\varepsilon$  and  $b^*$ . In the Japanese case, the electoral outcome in the proportional tier provides a good measure of the party support in the district. The empirical evidence indicates that the SDPJ is more likely to nominate candidates if the party has more support in those districts (Maeda 2008). This is because the SDPJ might nominate more aggressively in plurality district if necessary and the party is less likely to fail in recruiting the candidate ( $\varepsilon$  is smaller). Moreover,  $b^*$  will be smaller if the SDPJ has more support in the districts. If the SDPJ is more popular, its nomination is more harmful to the chances for DPJ victory ( $b_N$  is smaller), which in turn makes  $b^*$  ( $= (b_A - b_L) / (b_A - b_N)$ ) smaller. Hence, if the SDPJ has a stronger support base in the districts, the DPJ is more likely choose the leftist candidates. Therefore:

*HD1: The DPJ is more likely to choose leftist candidates if the SDPJ has more support in the district – more PR vote-share in the previous election given that the other conditions are held constant.*

However, as I have discussed in the extension, the DPJ might be less likely to change position according the perceived risk of the SDPJ's nomination, if the party has an incumbent candidate. In other words, the DPJ's position should be better adjusted to the SDPJ's strength in an open district than in the district with an incumbent candidate. Hence:

*HD1a: The DPJ is more likely to choose a leftist candidate if the SDPJ has more support in the district and the DPJ nominates a new candidate to the district, given the other conditions are held constant.*

The next section describes the statistical models that test these hypotheses.

### Statistical Models

The hypotheses of the chapter are classified into two groups: one is for the SDPJ's strategy over nominations measured with a dichotomous scale, and the other is for the DPJ's decision over candidate position measured with a continuous scale. I use a logistic regression for the SDPJ's strategy and an *OLS* for the DPJ's decision.

I use three models to test each hypothesis regarding the SDPJ's strategy. All of these models have the SDPJ's nomination in 2005 as the dichotomous outcome variable. The first model uses only the DPJ's position in 2003 as the explanatory variable to test *HS1*<sup>25</sup>. The second model adds the competitiveness of the district. As the indicator of competitiveness, I use the vote-share of the LDP out of the two major parties' (the LDP and the DPJ) votes. I define the competitiveness as the absolute distance of the vote-share from the tie ( $|\text{the LDP vote-share} - .50|$ ). Therefore, the term tests the validity of *HS2*. The larger the term is, the less competitive the district is. Finally, the third model includes the interaction term of competitiveness and the DPJ's position: the interaction term tests *HS2a*: if the SDPJ nominates their candidate when the DPJ's candidate is a centrist and the electoral outcome is predictable. On the other hand, the term for competitiveness in the model shows the effect of competitiveness on the SDPJ's nomination, given that the DPJ's position is equal to zero. As control variables, I use the

---

<sup>25</sup> As the explanatory variable for the strategy of the SDPJ in 2005, there are two available sources of the DPJ's candidate positions: their positions in 2003 and 2005. For the analysis I use the data of 2003 to examine the SDPJ's strategy in 2005, because the SDPJ probably did not have enough time to evaluate the DPJ's new candidate issue position in 2005. Few politicians expected the timing of the election in 2005, because the Prime Minister Koizumi suddenly dissolved the HR because his postal privatization proposal was turned down in the HC. Moreover, since only two years had passed since the last 2003 election, the DPJ had not finished candidate selection in many districts, and consequently rushed to choose candidates in those districts after the dissolution of the HR. Hence, it is unreasonable to suppose that the SDPJ had enough time to evaluate the new position of the DPJ's candidate, and reflect that information in designing their campaign strategy in 2005. Instead, the party is expected to rely more on the information that it already had in hand, even if the data could be outdated due to possible shifts in the incumbent position or the replacement of candidates. Therefore, in this chapter I use the DPJ's candidate position in 2003 to explain the strategy of the SDPJ in 2005.

SDPJ's PR vote-share and the presence of the SDPJ's candidates in 2003, because they are known to motivate small parties to run in plurality tier, even if they do not have a chance to win (Maeda 2008). Moreover, this term also controls the probability of the recruitment failure  $\epsilon$ .

Next, I use two models to test the hypotheses regarding the DPJ's positioning strategy. The outcome variable is the DPJ's candidate position in 2005, measured on a continuous scale. The first model uses the SDPJ's PR vote-share in 2003 as the explanatory variable of the model. The coefficient should be negative if the evidence supports *HD1*. However, the relationship between the SDPJ's support level and the DPJ's positioning might be strong because of the anchoring effect of the incumbency. Therefore, the second model uses the interaction term of the SDPJ's PR vote-share in 2003 and a new DPJ candidate in 2005. The interaction term tests *HD1a*: the DPJ runs the leftist if the SDPJ is strong in the district and the DPJ has a chance to choose a new candidate.

As control variables, I use the DPJ's ideological position in 2003<sup>26</sup>, the new candidate assignment in 2005, and the interaction of the two variables in both models. The control of the previous position has theoretical importance, because the coefficient of the explanatory variable shows the DPJ's effort to move to a point closer to the equilibrium, instead of the existence of the equilibrium per se. The DPJ's candidates should have a similar attitude between two campaigns if the same candidate keeps running; therefore the first term should have a strong positive effect. On the other hand, the interaction term should have a negative effect, because the candidate preference should be less similar if the DPJ replaced the candidate. Moreover, I control for competitiveness. If the race is expected to be close, based on the previous election, the DPJ may have an incentive to choose a centrist, because even small differences in votes owing to candidate positioning could be critical to the electoral outcome:  $b_A - b_L$  is large

---

<sup>26</sup> I use the DPJ's candidate position in 2003 and 2005 for the analysis as the control and the outcome variables. Therefore, the districts included to the analysis are limited to those that the DPJ nominated in both elections.

and consequently  $b^*$  becomes larger. I use the same competitiveness indicator with the SDPJ's model. Finally, I include the PR vote-share of the LDP in 2003 as a control for district preference: if the district is conservative, the DPJ has an incentive to nominate the hawk. Because there was no significant political party on the right side of the LDP, the LDP PR vote-share could be an indicator of the conservativeness of the district.

### Result

The columns of Table 3.1 show the outcome of each model regarding the SDPJ's strategy. The result of the first model clearly supports *HS1*: the SDPJ is more likely to run candidates if the DPJ is more hawkish, even if the SDPJ's PR vote-share and the presence of the SDPJ's candidate in the previous election are controlled. By contrast, the district competitiveness does not explain the SDPJ's decision well. The second column demonstrates that the competitiveness in the last election has insignificant effect on the SDPJ's decision in 2005. Moreover, the interaction term of competitiveness and the DPJ's candidate position in the third column does not have significant impact on the SDPJ's nomination. Although *Akaike Information Criteria (AIC)* suggests that those explanatory variables slightly improve the models' fit (the AICs are improved from -123.9 to -124.5 and -126.1 respectively), it can be concluded that there is insufficient evidence to support the hypotheses regarding competitiveness and the SDPJ's nomination: *HS2* and *HS2a*. The DPJ candidates' position has significant effect in the latter two models too, which indicates the robustness of the effect.

*Statistical significance* does not necessarily mean *substantive significance* of the effect. Therefore, I conduct a simulation for the counterfactual risk of the SDPJ's nomination given the DPJ's position and the SDPJ's support level, based on the estimates from the third model. The simulation shows that the dovish DPJ (positioned in the first quartile) has a SDPJ's opponent only with 2% probability, but the hawkish one (positioned in the third quartile) has 6% probability, given that the SDPJ had median levels of support. If the SDPJ's support was relatively strong, on the other hand, the SDPJ would nominate a candidate with 9% probability in response to the hawkish, and 22%

against the most hawkish candidate of the DPJ. For the DPJ, these numbers are not huge but hardly ignorable if the party plans their campaign strategy where the SDPJ have considerable support base.

Table 3.2 shows the result of the DPJ's positioning models. The first model suggests that the relationship between the SDPJ's support and the shift of the DPJ's position is not strong, although the direction goes along with *HD1*. By contrast, the second model clearly supports *HD1a*: the DPJ chooses the type of new candidates according to the strength of the SDPJ. On the other hand, the SDPJ's PR vote-share term without interaction has an even smaller effect in the second model. These results imply that if and only if the party chooses a new candidate, the DPJ moves the candidate position according to the SDPJ's support. Most of the control variables show the expected effect. The DPJ candidate positions are quite similar between the two consecutive elections. Unsurprisingly, the DPJ candidates have more similar preference in two elections if the same candidates continue running. Moreover, the more competitive the districts, the more centrist the DPJ's candidates become. However, the LDP PR vote-share -- the indicator of the district preference -- does not have a significant effect on the DPJ's position.

The simulation based on the estimates implies that with an additional 2.4% of the SDPJ's PR vote-share, the DPJ chooses new candidates with the median preference of the party, whereas the party would have nominated the hawkish otherwise: a candidate located in the third quartile. The size of the effect is not large given the strength of the SDPJ -- the median SDPJ's PR share was 4.2% in 2003 -- but still not ignorable. The DPJ is clearly attentive to the SDPJ's electoral base when it selects new candidates. All control variables have the expected effect except for the LDP PR vote-share, probably because the LDP PR vote-share is not a good indicator of the electorate's preference of foreign/defense policy.

## DISCUSSION AND CONCLUSION

This work investigates the *practical* logic behind the nomination strategy of a small radical party under the first-past-the-post electoral rule with two major centrist parties. Small radical parties, like the SDPJ in Japan, choose to campaign alone in a considerable number of districts despite the fact that the party has little chance of electing their candidates. They are unlikely to win seats, votes, or office with the strategy, at least in the short term. Moreover, these nominations might hurt their policy interests by contributing to electoral losses in the ideologically-closer major party. Hence, previous studies argue that the strategy is *ideologically* and/or long-term oriented, not *practically* oriented.

I demonstrate that this is not necessarily the case. The model in this chapter shows the *practical* logic behind small parties campaigning alone: by using the threat of nominations, they aim to achieve policy concessions from major centrist parties like the DPJ, which are afraid that the small parties' nomination would damage their campaign effort against major opponents. Contrary to the model by Tsebelis (1990), the small party does not actually have to be *ideological*. If the major party is not sure about the small parties' type, the *practical* small party also runs candidates as the ideological type does to attain policy concessions in subsequent negotiations. Moreover, the model shows that the small party would have no less incentive to nominate in competitive districts, where a nomination damages the major parties' chance to defeat their main opponents. .

The model explains the actual behavior of the small leftist SDPJ and the major center-left DPJ in Japan well. As the model predicts, the SDPJ selectively ran against the centrist/hawk DPJ. Moreover, the SDPJ's decision was not influenced by district competitiveness, due to the complex cost and benefit structure of the nomination. Because of the SDPJ's limited resources, the risk of a SDPJ's nomination to the DPJ was almost zero as long as the DPJ chose the leftist. However, the risk jumped up if the DPJ's candidate was a hawk regardless of the competitiveness, especially if the SDPJ were popular in the district. Furthermore, although we cannot observe the DPJ's belief directly, the results show that DPJ was attentive to the SDPJ's risk of the nomination in the districts where the party chooses new candidates. Paradoxically, this outcome implies the

empirical validity of the Downsian model, because it indicates that the DPJ attempts to nominate centrists if there should be no new serious entrants after two major parties chose their positions, as the original model assumes (Grofman 2004).

As for the empirical analysis, this chapter uses the strategic negotiation in the plurality tier of the Japanese MMM system, not the pure single-member plurality system, due to data availability on the issue position of individual candidates and the small parties' popularity in each district measured by their PR vote-share in the last election. However, the implication of this study is generalized to the strategies of small parties with extreme preference in the pure single-member plurality system. As long as these parties are a threat to the efforts of major centrist parties to defeat their main rivals ( $b^* < 1/2$ ), and can nominate without failure ( $\varepsilon < 1/2$ ), they can achieve the policy concession from the major parties.

This chapter does not intend to argue that this nomination strategy is *always* the most efficient way for small parties to achieve their policy goals. They can achieve office, seats and/or policy goals even under the plurality electoral system with the strategies discussed in the existing literature, such as pre-electoral alliances with other parties, as Komei did in Japan. The model of the chapter also implies that the nomination strategy is effective under the specific assumptions. If the small parties expect that these conditions, especially the assumptions that  $b^* < 1/2$  is not satisfied, they would choose alternative strategies. Moreover, even if the conditions hold, benefits of the alternative strategies could be larger than that of the nomination strategy.

The shifts of the SDPJ and the JCP strategies in the Japanese case can be good examples to discuss when the small parties choose the nomination strategy. The SDPJ chose to form a pre-electoral alliance with the DPJ in the 2009 HR election, and consequently the two parties did not compete except in few districts. The model and the empirical analysis of this chapter suggest that the completion of the DPJ candidate assignment to most of the districts have influenced the SDPJ's decision. The DPJ has been the largest opposition since 1997. Nevertheless, the party had many plurality districts without its own candidates in the late 1990's. The analysis indicates that the DPJ



chooses the candidate position according to the support level of the SDPJ, if the party nominates a new candidate to the district. In other words, once the DPJ completes the candidate assignment to most of the plurality districts, the SDPJ's nomination threat is less effective. On the other hand, this anchoring effect of the incumbents implies that the SDPJ succeeded in leaving enduring effects on the DPJ's preference with the nomination strategy, because the left leaning candidates of the DPJ chosen under the SDPJ threat will stay in office for considerable period. I do not exclude the other factors, such as the de facto coalition between the DPJ and the SDPJ in the HC since 2007, and the foreseeable (and realized) DPJ victory in the 2009 HR election and the subsequent chance of the coalition government between them, may have shifted the priority of the SDPJ's strategy. Nevertheless, the decreasing number of the DPJ open seats might have significantly influenced the change of the SDPJ's strategy.

On the other hand, the JCP's case suggests the importance of the ideological distance between the small and major party for the nomination strategy. The JCP does not seem to use their nomination to attain policy concession from the other parties. The party nominated in almost all the districts until 2005, and kept nominating in most of the districts in 2009 though the party failed to elect any candidates from the plurality tier. The ideological distance of the JCP with the other parties, especially with the DPJ could have made not only the pre-electoral alliance, but also the nomination strategy infeasible. The model in the chapter implicitly assumes that the electorates consider the two parties as alternatives: if they have too different an issue position, the small party supporters may show up to the voting booth only if their favorite party nominates a candidate and otherwise abstain. Therefore, the ideologically "too-distinct" small party may not be able to utilize the nomination option because their nomination does not influence the chance of the major party to beat the rival:  $b_A - b_N$  becomes very small and hence  $b^*$  is too large, which make the major party chooses the centrist candidates regardless the risk of the radical parties' nomination.

In summary, this chapter suggests a logic of small party nomination under the plurality electoral system. This helps to understand the behavior of the parties not only

in Japan but also those in the other countries. This article shows a promising direction of future research on party strategy through the combination of formal modeling and empirical analysis. The simple model in this chapter will explain the strategic interactions between parties in the other democratic countries too. Further investigation of the negotiation process will enhance our understanding of parties and representation, and how heterogeneous voice in democratic society is transmitted into policy outcomes.

## REFERENCE

- Chhibber, Pradeep K., and Ken Kollman. 2004. *The Formation of National Party Systems: Federalism and Party Competition in Canada, Great Britain, India, and the United States*. Princeton, N.J.: Princeton University Press.
- Cox, Gary W. 1997. *Making Votes Count: Strategic Coordination in the World's Electoral Systems*. Cambridge: Cambridge University Press.
- Cox, Karen E., and Leonard J. Schoppa. 2002. "Interaction Effects in Mixed-Member Electoral Systems: Theory and Evidence from Germany, Japan, and Italy." *Comparative Political Studies* 35 (9): 1027–1053.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper & Row.
- Duverger, Maurice. 1959. *Political Parties, Their Organization and Activity in the Modern State*. Array London: Methuen.
- Ferrara, Federico, and Erik S. Herron. 2005. "Going It Alone? Strategic Entry under Mixed Electoral Rules." *American Journal of Political Science* 49 (1):16–31.
- Ferrara, Frederic, Erik S. Herron, and Misa Nishikawa. 2005. *Mixed Electoral Systems: Contamination and Its Consequences*. Palgrave Macmillan, New York. N.Y.
- Gallagher, Michael and Paul Mitchell (eds). 2005. *The Politics of Electoral Systems*. Oxford: Oxford University Press.
- Golder, Sona Nadenichek. 2006. "Pre-Electoral Coalition Formation in Parliamentary Democracies." *British Journal of Political Science* 36 (2): 193-212.
- Grofman, Bernard. 2004. "Downs and Two-Party Convergence." *Annual Review of Political Science* 7: 25-46.
- Grofman, Bernard and Arendt Lijphart (eds). 1986. *Electoral Laws and Their Political Consequences*. New York: Agathon Press.
- Herron, Erik S. and Misa Nishikawa. 2001. "Contamination Effects and the Number of Parties in Mixed- Superposition Electoral Systems." *Electoral Studies* 20 (1): 63–86.
- Honaker, James , Gary King, and Matthew Blackwell. 2009. *Amelia II: A Program for Missing Data ver. 1.2-12*. Available from <http://gking.harvard.edu/amelia>.

- Horiuchi, Yusaku, and Natoru Ryota. 2007. "Nidai Seitōsei wo Sogai Suru Chihō Senkyo Seido" [The Impact of Subnational Electoral Systems on the Duvergerian Equilibrium: The Case of Japan] *Shakai Kagaku Kenkyu* [The Journal of Social Science]. 58 (5-6): 21-32.
- Imai, Kosuke, Gary King, and Olivia Lau. 2009. *Zelig: Everyone's Statistical Software ver.3.4-6*. Available from <http://gking.harvard.edu/zelig/>
- Kabashima, Ikuo, and Yoshihiko Takenaka. 1996. *Gendai Nihonjin No Ideorogī* [Ideology of the Contemporary Japanese]. Tokyo: University of Tokyo Press.
- Kabashima, Ikuo, and Koji Yamamoto. 2005. "The 2003-04 University of Tokyo/Asahi Shimbun Politicians' Survey Codebook." *Japanese Political Studies* 2(2):190-209
- Kabashima, Ikuo, and Koji Yamamoto. 2008. "The 2005 University of Tokyo/Asahi Shimbun Politicians' Survey Codebook." *Japanese Political Studies* 5 (1-2): 283-297.
- Karp, Jeffrey A., Jack Vowles, Susan Banducci, and Todd Donovan. 2002. "Strategic Voting, Party Activity, and Candidate Effects: Testing Explanations for Split Ticket Voting in New Zealand's New Mixed System." *Electoral Studies* 21(1):1-22.
- King, Gary, James Honaker, Anne Joseph, and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review*, 95 (1): 49-69.
- Kreps, David M, and Robert Wilson. 1982. "Reputation and Imperfect Information." *Journal of Economic Theory* 27 (2): 253-279.
- Laver, Michael, and Kenneth Benoit. 2006. "Estimating Party Positions: Japan in Comparative Context." *Japanese Political Studies* 3 (1): 109-133.
- Maeda, Ko. 2008. "Re-examining the Contamination Effect of Japan's Mixed Electoral System Using the Treatment-Effects Model." *Electoral Studies* 27(4): 723-731
- Müller, Wolfgang C. and Kaare Strøm (eds). 1999. *Policy, Office, Or Votes?: How Political Parties In Western Europe Make Hard Decisions*. Cambridge: Cambridge University Press.
- Palfrey, Thomas R. 1984. "Spatial Equilibrium with Entry." *The Review of Economic Studies* 51(1): 139-156.
- Reed, Steven R. 2001. "Duverger's Law Is Working in Italy." *Comparative Political Studies* 34(3): 312-27.

- Reed, Steven R. 2003. "Heiritsusei ni okeru shousenkyoku kouhosha no hireidaihyou tokuhyouritsu eno eikyou" [The Effect of the PR Tier on SMD Nomination Strategies in Mixed Member Systems: Some Evidence from Japan]. *Senkyo Kenkyu* [Electoral Studies] 18: 5–11.
- Reed, Steven R. 2007. "Duverger's Law is Working in Japan." *Senkyo Kenkyu* [Electoral Studies] 22: 96-106.
- Selten, Reinhard. 1978. "The Chain-Store Paradox." *Theory and Decision* 9: 127-259.
- Taniguchi, Masaki. 2006. "Syugiin giin no seisaku ichi" [Representative Positioning in Japanese Diet]. *Nihon Seiji Kenkyu* [Japanese Political Studies] 3 (1): 90-108.
- Tsebelis, George. 1990. *Nested Games: Rational Choice in Comparative Politics*. Berkeley: University of California Press.

## TABLES

Table 3.1: SDPJ Nomination in 2005

	Coeff. (Std.Err.)		Coeff. (Std.Err.)		Coeff. (Std.Err.)	
[Explanatory]						
D: Position 2003	1.02	(.422) *	1.49	(.557) *	1.48	(.698) *
Competitiveness 2003			4.05	(3.46)	6.12	(4.49)
Competitiveness 2003×D: Position 2003					-3.73	(5.30)
[Control]						
S: PR Share 2003	36.09	(13.05)**	36.56	(12.97) **	36.35	(12.87) **
S: Candidate 2003	1.53	(.555)**	1.49	(.557) **	1.48	(.558) **
(Intercept)	-4.98	(.775)***	-5.34	(.855) ***	-5.55	(.928) ***
n	242		242		242	
AIC	-123.9		-124.5		-126.1	

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

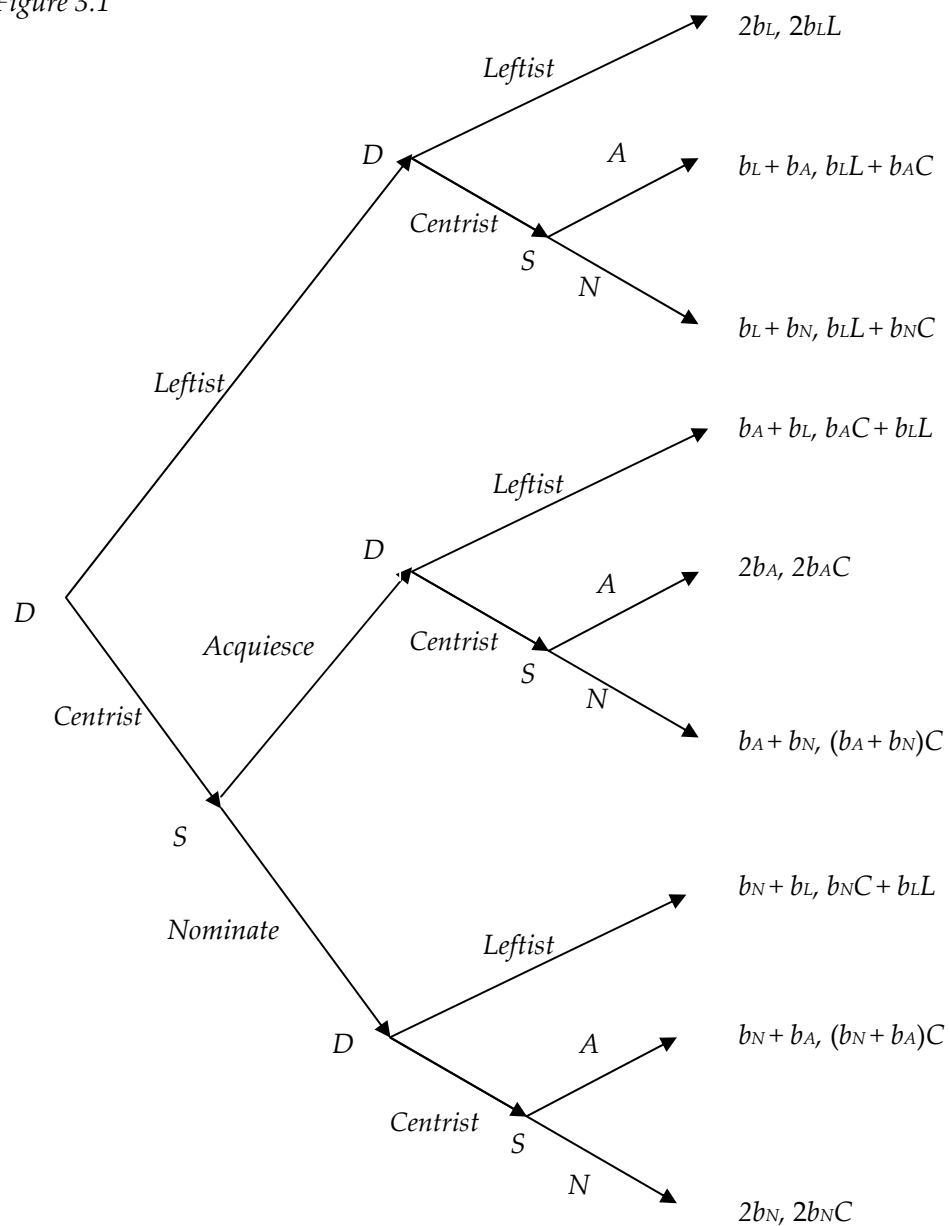
Table 3.2: DPJ Candidate Position in 2005

	<i>Coeff.</i> ( <i>Std.Err.</i> )	<i>Coeff.</i> ( <i>Std.Err.</i> )
<i>[Explanatory]</i>		
<i>S:PR Vote Share 2003</i>	-.998 (1.894)	-.300 (1.960)
<i>S:PR Vote Share 2003 × D: New Candidate</i>		-13.632 (6.361) *
<i>[Control]</i>		
<i>D:Position 2003</i>	.727 (.051) ***	.729 (.051) ***
<i>D:New Candidate</i>	.393 (.090) ***	.988 (.294) ***
<i>D:Position 2003 × D: New Candidate</i>	-.529 (.121) ***	-.560 (.121) ***
<i>Competitiveness 2003</i>	-1.323 (.549) *	-1.364 (.546) *
<i>LDP:PR Share 2003</i>	-.393 (.688)	-.308 (.683)
<i>(Intercept)</i>	.252 (.248)	.166 (.249)
<i>n</i>	242	242
<i>Adjusted R<sup>2</sup></i>	.475	.487

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

# FIGURE

Figure 3.1





## APPENDIX: WORDING OF THE QUESTIONNAIRES

We would like to ask your opinion to the constitution:

1) The Amendment of the Constitution: Should we amend the constitution?

For the following opinion, do you agree or disagree?

2) The Size of Defense Force: Japanese defense force should be reinforced.

3) The Japan-US Alliance: The Japan-US alliance should be reinforced.

4) The Preemptive Strike: If the other countries are expected to attack us, we should not hesitate to conduct a preemptive strike.

5) Seeking the Permanent Seat of the UN Security Council: Japan should seek the permanent seat of the UN Security Council to play more active international role.

6) The Foreign Policy to North Korea: We should give more weight to pressure rather than dialogue with North Korea, to drive the country into the regime change.

7) The Right of the Collective Defense: The Japanese government should change the current interpretation of the constitution to exercise the right of collective defense.

## **Chapter 4**

### **Challenging Strongmen with Zombies: Incumbency Advantage and the Consequence of the Double Candidacy System**

#### **INTRODUCTION**

The quality of democracy is fundamentally dependent on democratic responsiveness -- how much the political process induces the government to form and implement policies that the citizens want (Powell 2000). The electoral process is one of the most important factors in assuring this responsiveness. However, if the process advantages incumbent representatives or parties that held incumbency over non-incumbent opponents, democratic responsiveness deteriorates because unpopular incumbent representatives/parties may stay in power, due to their incumbency status. Poorly performing governing parties may not be punished if parties obtain extra electoral gains from fielding unpopular incumbent candidates, which offset the decline in party votes. Similarly, the electorate finds it difficult to remove unpopular incumbents from office (Ariga 2010).

Given the significance of these normative concerns, a number of scholars have studied incumbency advantage in democratic countries. Owing to the very high re-election rate in the US House of Representatives in the last few decades, several studies examine incumbency advantage in American politics (e.g. Ansolabehere and Snyder 2002; Carson et al. 2007; Cox and Katz 2002; Erikson 1971; Fiorina 1977; Gelman and King 1990; Mayhew 1974). Studies of incumbency advantage in other countries have been relatively few and are concentrated in the countries with the single member

plurality electoral system like that in Britain (Gains 1999; Katz and King 1999; Linden 2004; Uppal 2009). However, some scholars have recently started to examine the magnitude of the advantage under different electoral systems (Ariga 2010; Hainmueller and Kern 2008; Hirano and Snyder 2009; Titunik 2008).

Since 1994, the Japanese House of Representatives (HR) has used the mixed-member majoritarian electoral system, which combines single-member plurality and proportional representation (PR) systems as two tiers of a single electoral system. This chapter measures the incumbency advantage in plurality contests under the mixed system. This is not a straight forward problem, however. I show that *Regression Discontinuity Design (RDD)*, a method recently applied in political science to estimate the advantage under plurality electoral systems (e.g. Ariga 2010; Hainmueller and Kern 2008; Lee 2008; Linden 2004) does not work in the Japanese case. The double candidacy system – a seemingly minor electoral rule in Japanese electoral system – enables parties to provide their PR seats to the most competitive losers of the plurality contests. Because of this double candidacy, *RDD* cannot measure the electoral advantage of the district incumbents over the challengers without seats, as designed. Instead, I utilize another mechanism of the system to estimate the advantage with *a natural experiment*.

This apparently technical topic, on the other hand, implies an unexpected contribution of the double candidacy system to competitiveness, and subsequently democratic responsiveness in Japanese elections. The double candidacy system has been criticized as helping unpopular incumbents retain their seats, even if they lost in plurality contests (e.g. McKean and Scheiner 2000). However, in this chapter I demonstrate that the system actually hurts the incumbent by awarding available PR seats to the most competitive candidates in each party who failed to win at the plurality district-level. In the next election, these representatives challenge at the district-level with incumbency status from the PR seats, canceling out the electoral advantage of the district-level incumbents. Hence, the system contributes to competitive, and consequently more responsive, elections in Japan. Moreover, this mechanism also helped the Democratic Party of Japan (DPJ), the current governing party, become a serious

contender to the Liberal Democratic Party (LDP). The LDP has been dominant in Japanese politics for more than half a century, based on a strong support base in rural areas. I demonstrate that without the double candidacy system, the LDP could have kept their dominant status for a much longer period, due to an initial electoral advantage through controlling more incumbent seats, especially in rural districts.

This chapter is organized as follows: In the next section, I describe the background of this chapter. The section starts with the previous discussion regarding incumbency advantage, especially its various sources and influence upon democratic responsiveness. Next, I briefly explain the current electoral system of the Japanese HR, the mixed-member majoritarian electoral system, by focusing the double candidacy system. In the following sections, I describe the data and the statistical approaches, and the results of the analyses for each. The final section discusses the substantive implication of the results, in particular the role of the double candidacy system, in facilitating more responsive elections in Japan, and consequently the DPJ's challenge to once-dominant LDP.

## **BACKGROUND**

This chapter investigates *incumbency advantage*: the electoral advantage of the candidates holding incumbency over opponents without seats. Scholars have discussed various sources of the advantage. The classification of its sources – quality of the candidates, personal office holding and party incumbency – helps us to understand its potential effect on democratic responsiveness.

Levitt and Wolfram discuss first two distinct sources of incumbency advantage: the quality gap between incumbents and challengers, and direct benefits to office-holders (Levitt and Wolfram 1997). Incumbent candidates in general are more experienced and attractive than average candidates. Otherwise, they would not have been elected in the previous election. Next, incumbent candidates tend to have more electoral resources than their opponents, owing to their status as the office holder. Incumbents are able to establish personal ties with their constituents more efficiently by

lobbying for local interests in the legislating process, as well as by providing pork in the form of local public projects. Moreover, mass media covers political activity of the incumbents more frequently than other (potential) candidates before the campaign. Furthermore, in many democratic countries, the incumbent representatives receive public funding for their political activities in addition to their salary (Levitt and Wolfram 1997).

Party incumbency can also help new candidates defeat their opponents, even if the incumbent from their own party retires. The local party organizations develop faster if the party has incumbency in the districts since, like individual incumbent representatives, parties establish personal ties with the voters more efficiently. Moreover, the incumbents in the legislature can help candidates of the same party run in the other levels of the election, such as the local mayoral election. This, in turn, helps not only the candidates themselves but also the other candidates who run from the same party in the area.

If the incumbents have an electoral advantage due to their superior quality, the advantage does not harm democratic responsiveness: the electorate votes for more qualified candidates, who happen to be incumbents. However, if the individual candidates and parties have an electoral advantage solely because of holding office, it can hurt democratic responsiveness. Unpopular incumbents still win races by utilizing resources attained through their office. Similarly, governing parties may not be punished, even if they perform poorly, if the parties have electoral advantage from simultaneously fielding several incumbent candidates. Moreover, representatives may not seriously pursue policies reflecting the preference of their constituents if they expect that they can win the next election with certainty (Griffin 2006).

Furthermore, a large incumbency advantage could drive out qualified challengers, such as those working as public officials in a different level of government from the contest, due to their lower chance of attaining office and high opportunity cost. This could lead to a self-reinforcing incumbent dominance. Regardless of the short-term rise and decline in popularity, the incumbents continue to win because everybody

expects they win, which in turn helps the expectation to be realized. Scheiner (2005) argues that this self-reinforcing mechanism explains the LDP dominance in the Japanese rural districts.

The magnitude of the incumbency advantage, on the other hand, implies the pattern of electoral competition in countries. If electorates mainly care about national political issues to decide their voting, such as the governing party's management of the macro economy and/or policy programs that parties promise to pursue, they care little who/which party occupies the incumbency in their districts. If the voters place a high value on incumbency, in contrast, it suggests that personalistic factors and/or narrowly focused interests play an important role in the campaign.

Japanese campaigns have been often described as pork-oriented and personalistic, rather than programmatic and party-centered. Many scholars attribute this tendency to the multi-member district/single non-transferable vote electoral system (MMD/SNTV) of the HR used from 1947 to 1994 (e.g. Curtis 1971; Kohno 1997; Ramseyer and Rosenbluth 1993). If a party attempted to achieve a legislative majority under the system, the party had to nominate multiple candidates per district. These candidates, in turn, were motivated to differentiate themselves with other co-partisan candidates not on the overlapping policy program, but on personalistic achievements, such as pork-provision to their particular constituents.

Due to money politics and clientelism endemic to personalistic campaigns under the MMD/SNTV, in 1994 Japan changed the HR electoral system to a mixed-member majoritarian electoral system to change the pattern of electoral competition. The new electoral system combines two traditional electoral systems, the single member plurality and PR system, as two electoral tiers of a single electoral system. The plurality tier has 300 single member districts, and the PR tier has 11 regional PR districts with 200 (later reduced to 180) seats.

The Japanese electorate has experienced five general elections under the system (1996, 2000, 2003, 2005 and 2009). So far, the evidence on how the electoral reform has changed the electoral competition in Japan is mixed (e.g. Christensen 1998; Taniguchi

2004). There was significant party realignment in the 1990s. However, the party system was gradually transformed into a two party system under the new electoral system, dominated by the LDP and the DPJ (Reed 2001). The LDP has been predominant in Japanese politics for more than fifty years. Although the party was temporally out of office after a party breakup in 1993, it quickly returned to power in 1994 after the electoral reform legislation. The LDP succeeded in keeping its power base, in particular by regaining its dominant status in rural areas, and stayed in office as the coalition government until 2009. On the other hand, the DPJ was formed in 1996 by the center-left politicians as the second-largest opposition party at that time. The party became the largest opposition in 1998, and grew to be the unique and serious contender to the LDP rule. Although the DPJ lost badly in 2005, owing to Prime Minister Koizumi's popularity among the urban electorates – which the DPJ had been relying on – the DPJ made a sweeping victory in the 2009 HR election and seized power from the LDP.

The DPJ's victory suggests a transformation of electoral campaigns in Japan, because the LDP had been predominant in Japanese politics based on its organizational grip of rural voters via candidates' personal networks. For example, Reed, Scheiner and Thies (2009) discuss that party affiliation became a stronger predictor of electoral victory than individual candidate attributes in the 2009 HR election. However, the DPJ seems to not be immune to pork-oriented politics, according to the scandals of the former party leader Ichiro Ozawa. By measuring the magnitude of incumbency advantage, I attempt to evaluate the current pattern of electoral competition in Japan.

In this chapter, I demonstrate that the double candidacy system -- a seemingly minor electoral rule in Japanese electoral system – is very important, both for technical and substantive reasons: the accurate estimation of incumbency advantage on the one hand, and its impact to electoral competitiveness and democratic responsiveness on the other. The mixed-member system in Japan allows the parties to nominate their plurality candidates to the PR list as well. If these double candidates win in plurality contests, their names are removed from the PR list. If they lose, they still have a chance to be elected from the PR tier as “Zombies”: *dead* in plurality district elections but *revived* at

the PR level, if their ranks are high enough on the list. Moreover, the parties are able to list these double candidates at the same rank on the PR list. Among the losing candidates listed in the same rank, those achieving the highest vote-share against their respective winner are given priority. Because of the rule, candidates who lose by a narrow margin in plurality districts almost surely receive PR seats, and usually challenge the district winners in the next election.

The system functions as follows: Suppose that Party X nominates five district candidates A, B, C, D and E to the first rank of PR list as double candidates. Candidate A won in the districts and consequently is removed from the PR list, and the other four lost but achieved 85%, 95%, 80% and 90% of the district winners' vote-share, respectively. If Party X achieves votes enough to elect two candidates from the PR tier, PR seats are allocated to candidate C and E, because they attained a better vote-share ratio than candidate B and D, who are nominated at the same rank on the PR list.

Due to the mechanisms of the double candidacy system, it is technically difficult to estimate the incumbency advantage with a *Regression Discontinuity Design*, used in previous studies of plurality electoral systems (e.g. Ariga 2010; Hainmueller and Kern 2008; Lee 2008; Linden 2004). Technical details will be explained in the following section, but *RDD* compares the shift in the vote-share between two political camps for two consecutive elections in very competitive districts. If a candidate who won with a very narrow margin and achieved incumbency in a plurality contest wins convincingly in the next election, it can be inferred that the incumbency of the candidate provides an electoral advantage against an opponent without seat. However, under the double candidacy system, the candidates who lost by a small margin usually gain PR seats, as candidates C and E in the example. Hence, in the very competitive districts that *RDD* uses to estimate the incumbency advantage, both camps hold incumbency in the next election -- one is from the plurality district election and the other is from the PR tier. Therefore, *RDD* measures not the incumbency advantage of the district incumbents to the challengers without seats, but that of the district incumbents to PR incumbents. In



the empirical section, I demonstrate, with *RDD*, that district incumbents do not have significant electoral advantage over PR incumbents.

To estimate the incumbency advantage, this chapter instead utilizes another mechanism of the double candidacy system. If the candidates belong to different regional PR districts, different incumbency status (PR or none) is conferred to candidates from districts with the same losing vote-share ratio.

Substantive implications of the double candidacy system are yet not well understood. Some scholars criticize the system because it helps unpopular incumbents retain seats (e.g. McKean and Scheiner 2000). This chapter, in contrast, argues that the system actually hurts incumbents. The system enables each party to award its PR seats to the most competitive losers. These *Zombies* challenge at the district level by utilizing their incumbency status from the PR seats, which cancels out the advantage of the district incumbents.

The mechanism contributes to electoral competitiveness in plurality contests, and consequently democratic responsiveness in Japanese elections. First, the system makes it difficult for the district incumbent candidates or the parties fielding the incumbents to monopolize incumbency status. If their opponents achieve a certain level of vote-share to the district winners, they attain seats from the PR tier. As criticized by previous studies, these *Zombie* incumbents often behave as if they are the representatives of the districts, although they are actually elected from the PR tier (McKean and Scheiner 2000). However, by acting for the local interest with the electoral resources available as the office holder, they can more easily establish a support base to challenge the district incumbents in the next election, which cancels out the electoral advantage of the district incumbents.

Moreover, the system helps to reduce quality gaps between incumbents and challengers by improving the latter's chance of gaining seats. Facing strong incumbents, challengers do not have to win their first election to gain seats. They are entitled to receive seats at least from the PR tier, as long as they achieve a certain vote-share. As a result, qualified challengers have more incentive to participate in campaigns against

strong incumbents. Hence, the double candidacy system may work against the electoral advantage of the district incumbents not only by canceling the office-holding advantage by assigning “Zombie” incumbency status to the best losers, but also by improving the quality of challengers. This point will be further elaborated with both quantitative and qualitative approaches.

## EMPIRICAL ANALYSIS

In this section, I conduct empirical analyses to estimate incumbency advantage in Japan. This section starts with a description of the data, by focusing on how the outcome variable and the main covariate are defined in this chapter. Next, I discuss two approaches to estimate the incumbency advantage: *RDD* and *a natural experiment*. Finally, the results of these analyses are presented.

### Data

For the empirical analysis, this chapter uses the electoral outcomes of plurality contest in the Japanese HR election. Therefore, the units of analysis are the single member districts of the HR. I use the electoral results in the 2005 election as the outcome variable, and those in 2003 as a main covariate.

I choose these elections to obtain reliable inference. Given the heterogeneity of party support across districts, it is helpful to have the vote-share of the parties within each district for the two consecutive elections. Unfortunately, the party system in Japan was unstable from 1993 to 2003. Moreover, redistricting in 2002 hinders comparison of the outcomes of many districts between the 2000 and 2003 HR elections. In other words, the 2005 HR election is the first election after the electoral reform for which we can derive reliable inference from the analysis with two consecutive elections.

For the analysis, the electoral result of each district is recoded as follows. Since 2003, there have been two major political “camps” in Japanese politics: the LDP and the

DPJ<sup>27</sup>. I use the vote-share of the LDP-camp candidates<sup>28</sup> as a proportion of the votes attained by both the LDP- and DPJ-camp candidates. Hence, if the LDP-led coalition candidate gains 50% of the total votes, the DPJ achieves 45%, and the other parties attain 5% in total<sup>29</sup>, the outcome variable is coded as .526 ( $=50/(50+45)$ ). Because only candidates from these two major camps have serious chance of victory in the plurality contests, the result indicates the victory of the LDP if it is more than .50, and that of the DPJ if less than .50. This number is a relatively reliable and stable measure of electoral power balance in the single member districts of Japan, especially if the candidates of the other parties did not achieve considerable vote-share<sup>30</sup>.

---

<sup>27</sup> There are other small parties in Japan that have the independent campaigns, such as the Japanese Communist Party (JCP) and the Socialist Democratic Party of Japan (SDPJ). Although they nominate their candidates not only to PR tier but also to plurality districts, they cannot elect them in the latter.

<sup>28</sup> Most of them were affiliated with the LDP, but few candidates from the LDP's coalition partner, Komei Party and the New Conservative Party (NCP) also ran from the camp. The NCP dissolved in 2003 after the election, and their MPs joined to the LDP.

<sup>29</sup> The other small parties, such as the SDPJ and the JCP achieved some amount of the votes in the plurality contests in the 2003 and 2005 elections. The effective number of candidates in the plurality campaigns was close to two for these two elections (around 2.3 in both elections). The JCP nominated in most of the districts (300 in 2003, 275 in 2005), in addition the SDPJ ran candidates in considerable number of the districts.

<sup>30</sup> I exclude some districts from the analyses because of this reason. First, I did not use the districts where the conservative camp split up. Owing to the personalistic nature of the political organization, the LDP affiliated politicians often run as independents if they cannot attain the party nomination, but they usually join to the LDP if they win. The participation of these independent, in turn, radically changes the LDP vote-share even in the consecutive elections. Moreover, the LDP had many defectors in the 2005 HR election, which left the party because of the inter-party conflict over the postal service privatization. The contests between these ex-LDP defectors and the LDP *assassins* – the LDP new candidates nominated to oust the defectors also make it difficult to compare the electoral results for two consecutive elections. Therefore, I exclude the districts in which more than two candidates run from the conservative camp. In addition, I also exclude the districts in which the DPJ did not nominate the party candidates, and those in which the DPJ failed to obtain the largest vote-share among the opposition camp. The list of these districts is available from the author by request.

## Method

To measure the incumbency advantage and the function of the double candidacy system in Japan, I use two approaches, *RDD* and *a natural experiment*. First, I use *RDD* to estimate incumbency advantage in plurality contests for the HR. However, as described in the previous section, this method is expected to produce biased results in the Japanese case because it measures the electoral advantage of the district incumbents over the PR incumbents, not district incumbents to new challengers, as designed, because of the double candidacy system. I then explain this point in more detail through an explanation of the theoretical mechanisms of *RDD*. Next, I leverage a natural experiment to estimate incumbency advantage by utilizing another mechanism of the double candidacy system. This approach estimates the electoral advantage of PR incumbents over challengers without seats, as opposed to the list incumbents of another party.

### **[The Incumbency Advantage with the *Regression Discontinuity Design*]**

Recently, scholars have applied *RDD* to estimate the incumbency advantage (e.g. Ariga 2010; Hainmueller and Kern 2008; Lee 2008; Linden 2004). Traditionally, the technical hurdle of measuring the incumbency advantage has been the extrapolation problem (King and Zeng 2007). Usually, there is no overlap between the districts that a party won in the previous election (to achieve incumbency) and those that the party did not. If a party won in the previous election, it means that the party attained more votes than any other party candidates in the election. On the other hand, if the party lost, it means that another party or independent candidate gained more vote than the party candidate. It is generally impossible to both win and lose the same district simultaneously. Moreover, the districts which a party won and those which the party lost are different not only in the party's vote-share, but also in other, often unobservable, attributes, such as the quality of candidates.

*RDD* uses the quasi-experimental setting to avoid the extrapolation problem. The design is used when there is a clear cut-off point according to a value of the covariate in the treatment assignment: The units receive one treatment if the covariate  $X$

$\geq c$  when  $c$  is the cutting point, but the units receive another if  $X < c$ . The important point is that the treated and the controlled units do not share the same value of the covariate  $X$ . The value of  $X$  is always larger (or smaller) than  $c$  for the treated units, but smaller (or larger) than  $c$  for the controlled units.

The design utilizes differences in the units within a small range surrounding the cutting point. Within the range, the treatment is expected to be assigned due to a small, random fluctuation of the covariate. As a result, these units should have almost the same distribution of the covariate, regardless of assignment. If there is any difference in the outcome between observations on either side of the cutting point, this is attributable to the difference in treatment. Hahn, Todd and Van der Klaauw (2001) and Porter (2003) provide a formal theoretical framework and estimation methods.

In the context of incumbency advantage, suppose that a country has the single member plurality electoral system and two major parties, Party A and B. A district in which party A achieves 50.1% of the vote-share must be very similar to the district where party A acquires 49.9% of the total votes. However, the electoral rule assigns different incumbent candidates/parties to these districts based on small differences in vote-share. In the former case, Party A gains the incumbency but Party B does not, but in the latter case the roles are reversed. These districts should be quite similar except their incumbency status. Therefore, the difference in their vote-share at the next election can be ascribed to their different incumbency statuses.

The incumbency advantage measured by *RDD* is designed to derive from party incumbency and the office holding if the incumbents keep running. This design does not measure the advantage owing to the candidate quality, because the quality of the candidates is assumed to be balanced between the two treatment groups within the small range across the cutting point. Moreover, the method does not measure the incumbency advantage in general but the advantage in the very competitive districts -- the estimate is the local average treatment effect, the *LATE* (Imbens and Angrist 1994). However, usually the advantage in the competitive districts is expected to be more

important than those in the other districts, because it has larger influence on the party's seat share.

Nevertheless, because *RDD* depends on the observations from competitive districts, it does not to work well in the Japanese case, at least in measuring the incumbency advantage of the district incumbents over challengers without seats as originally designed. As explained, the double candidacy system allows parties to assign their PR seats to the most competitive losers of the plurality contests. In other words, as long as parties use PR seats for this purpose, the district incumbents always have to confront opponents with PR incumbency in the most-competitive districts. The cutting point at an LDP vote-share of .50 does not shift the district treatment status from “the DPJ achieves incumbency but the LDP does not” to “the LDP achieves incumbency but the DPJ does not”, but rather from “the DPJ gains district incumbency but the LDP has PR incumbency” to “the LDP attains district incumbency but the DPJ instead achieves PR incumbency”. This chapter conducts the analysis with *RDD* as previous studies did. However, I expect that *RDD* does not detect the incumbency advantage at the cutting point, unless Japanese voters consider district and PR incumbents differently.

To estimate the local average of the outcomes from both sides of the cutting point, I use the kernel weighted local polynomial smoothing line with triangle kernel function. This method uses the non-parametric fitted lines that give more weight to the observations close to the estimating point. The estimate relies less on the parametric assumption or the covariates balance between the observations far away from the estimating point. I set the bandwidth of the smoothing lines to .0250. The criterion is subjective, however. The smaller the bandwidth is, the less biased the estimate will be but it will also be less efficient, because the non-parametric kernel function uses more nearby but fewer total observations for the estimation. To examine the robustness of the estimate for the bandwidth, I also estimate the treatment effect with 50% (bandwidth=.0125) and 200% (bandwidth = .0500) of the initial bandwidth .0250. I use the *Stata* ado-on module *RDD* by Nichols to estimate the treatment effect and the bootstrap standard errors of the estimate (Nicols 2007).

### **[The Zombie Incumbency Advantage with the Instrumental Variable]**

Next, I use a *natural experiment* to estimate the incumbency advantage in Japan, by utilizing the mechanism of the double candidacy system. This method measures how much more the candidates/parties with PR incumbency gain in electoral outcomes than those without, given the other conditions are held constant. It measures a different type of the advantage from that measured by *RDD*. However, if this method indicates that the PR incumbents achieve better electoral outcomes than the challengers without seats, and *RDD* in this chapter shows that the district incumbents do not have electoral advantage to the challengers with PR incumbency, these results in combination imply that the PR incumbency cancels out the electoral advantage that the district incumbents could have if they competed against challengers without seats.

As for district incumbency, the main problem in estimation of PR incumbency is lack of the overlap in the electoral outcomes from the previous election, between the districts that a candidate lost and achieved a PR seat on the one hand, and those where the candidate failed to acquire seats from both the plurality and PR tier on the other. I utilize the mechanism of the double candidacy system that requires different minimum vote-share for lost candidates to acquire a PR seat if the districts belong to different regional PR districts.

Take an example, suppose that Party X nominates candidate A to the district J1, in addition to listing her at the first rank of the region J's PR list, alongside the other district candidates in the region. Party X also nominates candidate B to the district K1, and list him at the first rank of the region K's PR list with the other plurality candidates. Both candidate A and B lose in their district competition but achieve 85% of the winner's vote-share. Although they achieve the same relative vote-share, they can receive different incumbency status in the next election, because they are nominated in different regional PR lists.

In the 2003 HR election, the DPJ double candidates who lost in the plurality race received PR seats if they achieved 67% of the LDP winners' vote-share in Hokuriku

region. In contrast, in the Tokyo region they were required to reach an 86% level to gain a PR seat. These were the extreme cases, and hence the DPJ double candidates succeeding in the range (the LDP's vote-share is between 53.8% and 59.8%) had non zero probability of receiving either of the incumbency status. Similarly, the minimum comparative vote-share of the LDP *Zombie* candidates ranged from 85% to 95% – the LDP's vote-share is between 46.0% and 48.9% -- among the LDP double candidates nominated on the same rank of the PR list<sup>31</sup>. In 2003, there are 43 districts that fell in the former range, and 26 for the latter.

This chapter measures the electoral advantage of the candidates with PR incumbency to those without seats by comparing the electoral outcomes of the districts within these ranges of vote-share in the subsequent election. There is a convenient attribute within Japanese elections which eases use of this research design for estimating the effect of PR incumbency. Japan is divided into 11 PR districts, and each region should not have any direct effect on the electoral outcome except via the allocation of the PR seats to the double candidates. In other words, the PR region can be used as the *instrumental variable*. Two districts in which the LDP gained the same vote-share asymptotically have the same distribution of both observed and unobserved covariates. Hence, given the LDP's vote-share in the last election, the difference in the electoral outcome between the districts that receive PR incumbency or none owing to PR region is attributable only to their incumbency status.

This method does not measure the electoral advantage of the PR incumbents in general, but only in the districts falling within the range (*LATE*). Therefore, the size of the effect could be different for very competitive (ex. the LDP vote-share is almost .50) or uncompetitive districts (ex. the LDP's vote-share is less than .40 or more than .60). It is

---

<sup>31</sup> I do not use the districts that the LDP double candidates are assured to attain PR seats even if they lost by a large margin, by ranking them not in the same rank with other double candidates but in the very high-rank of the PR lists. These districts had the Democrat cadres as the incumbents, and the LDP used these PR seats as an incentive for potential candidates to challenge them. Therefore, these districts have quite different attributes from the other districts. Moreover, the region does not assign them the *Zombie* status exogenously to these districts.



impossible to extend inference to them from the observations without taking risk of undue extrapolation.

I use *OLS* to estimate the advantage of Zombie incumbents over non-incumbent candidates. The explanatory variable is Zombie incumbency: the PR seats provided to the double candidates who lost in the plurality contests. Because the hypothesis is that a Zombie incumbent performs *better* than a challenger without a seat, the statistical test is single-sided. The outcome variable is the same with the *RDD* in the previous section: the vote-share of the LDP-led coalition as a fraction of the major camps vote-share in 2005.

The *OLS* model uses the vote-share in 2003 and its square as the main covariates. The quadratic term are used because both the LDP and the DPJ performed better in 2005 where they did not previously. In the 2005 HR election, Prime Minister Koizumi's proposal to reform the postal service attracted many urban electorates but antagonized the traditional LDP electoral base in rural areas, and consequently the LDP candidates did better in urban districts than rural ones. By the same token, the DPJ candidates fought better in rural districts, even though the DPJ had been a relatively urban party before 2005. Therefore, the vote-share in 2003 is expected to have a non-linear relationship with that in 2005. The quadratic term is used to control for the non-linear support shift in the election. I subtract .50 from the vote-share before taking the square. Moreover, I include the participation and exit of the SDPJ to the plurality contest to the model for the efficiency, because they should have considerable effect of the DPJ vote-share. The participation of the SDPJ, a small leftist party, should decrease the vote-share of the center-left DPJ in the districts, and their exit should increase the DPJ's vote-share. In contrast, I did not add the nomination of the JCP, another small leftist party to the model, because the party nominated their candidates in most of the districts by 2005.

## Result

Table 4.1 shows the estimation of the incumbency advantage with *RDD*. The estimate is almost zero with the bandwidth at .025, with .0125 (50% of the original) and .050 (200%

of the original). The results seem to be insensitive to the choice of bandwidth. All of these coefficients were statistically insignificant, based on boot-strapped standard errors. The result suggests that district incumbents do not have an electoral advantage from party incumbency and holding office in 2003 over challengers holding PR incumbency<sup>32</sup>.

Next, Table 4.2 shows the result of analysis regarding the effect of PR incumbency on the electoral outcome through the natural experiment. The result shows that both LDP- and DPJ-camp candidates performed better in 2005 in districts where their double candidates achieved PR seats in 2003, given that vote-share in 2003 and the SDPJ nominations are held constant. The PR incumbents of the LDP camps acquire more than 3% higher vote-share (within the two major political camps) than the same party challengers without seats. Similarly, DPJ PR-incumbents achieved 2% better vote-share balance<sup>33</sup>. Although the number of cases is relatively small (43 and 26), they are statistically significant at the 5% level.

## INFERENCE AND DISCUSSION

The analyses in this chapter shows two results: First, the district incumbents in Japanese HR do not have an electoral advantage over their opponents in subsequent elections, at least in the competitive districts. Next, PR incumbents perform better (2-3%) than the same parties' candidates without seats if fighting against the district incumbents of the other party, even if the party achieved the same vote-share in the last election. In

---

<sup>32</sup> I conduct the same analysis again without open districts. *RDD* usually estimates the electoral advantage owing to the office holding and the party incumbency. However, if the incumbents retire, the party has the advantage only owing to the party incumbency in the districts, even if the party won in the previous election. If many incumbents retire on the other hand, especially in competitive districts, it is not clear what the design measures. To measure the effect truly owing to the office holding and the party incumbency, I conduct the analysis without the open districts. I achieve almost the same result and the conclusion inferred from the analysis does not change, though there still remains the concern to strategic retirements (Ansolabehere and Snyder 2004).

<sup>33</sup> The negative coefficient means that the LDP-led coalition candidates achieve fewer votes in the districts that their DPJ opponents held PR incumbency, in other words, the DPJ PR incumbency increases the DPJ's vote-share.

combination, these results imply that district incumbents should have an electoral advantage over challengers, if the challengers do not hold PR seats. Once a party loses in a very tight “open” race, the party would have been disadvantaged in total around 5% (2+3) in the next election, in comparison to the case that the party won and achieved the incumbency.

In short, the double candidacy system contributes to more competitive campaigns in the single member districts of Japanese HR. If the double candidacy system had not existed, the district incumbents in Japan would have non-ignorable electoral advantage over challengers, as is the case in other countries. However, the district incumbents in Japan usually have to compete with challengers holding PR seats, at least in the competitive districts, who utilize their office to counter the status of the incumbents. Moreover, the system supposedly gives qualified candidates more incentive to challenge the incumbents by providing a better chance of achieving the seats, although this chapter cannot measure that effect of the system directly. Hence, counter to the conventional wisdom, it can be concluded that the system functions against the interest of the incumbents, and consequently leads to more competitive elections in Japan. The competitiveness in turn contributes to more democratic responsiveness in Japanese election, because the voters are able to replace unqualified incumbents more easily if the campaign is competitive.

Although the system itself works impartially, there are reasons to believe that the double candidacy system helped the DPJ, the current governing party, to become a serious contender to the LDP, which has been predominant in Japanese politics for more than a half century. First, the anti-incumbency effect of the system functions against the party with more incumbent representatives. In the late 1990s and the early 2000s, the LDP fielded many more incumbent candidates than the DPJ did, especially in rural areas where the party had established a strong support base. Without the system, the LDP could have kept its dominant status for a longer period by utilizing the advantage that these incumbent candidates had over their opponents.

Moreover, the LDP and the DPJ used the double candidacy system quite differently. Until recently, the LDP used PR tier to avoid intra-party conflicts and provided seats for the excessive number of incumbents, instead of saving the “best losers” of plurality contests. The electoral reform in 1994 corrected malapportionment between urban and rural districts during the MMD/SNTV era. Moreover, the reform reduced the number of representatives elected from the plurality districts from 512 to 300. Therefore, the LDP had more incumbents than the number of single member districts, especially in rural areas. To avoid the internal strife that could be fatal to the plurality electoral contests, the LDP persuaded these “excess” incumbents to run from the PR tier, by nominating them to the highest rank of the PR list (Niwa 1997). Furthermore, in some districts the PR seats are used as incentive for the ex-LDP politicians to return to the party without conflicts over the nomination (Kage 1997).

In contrast, the DPJ used PR tier primarily to save the best losers in the plurality contests with the double candidacy system, as discussed in this chapter. There is no reliable evidence to believe that this strategy is taken based on the knowledge to positive effect of the double candidacy<sup>34</sup>. However, in hindsight the decision helped the party to become a major contender to the LDP. The DPJ started as a medium size party in 1996. The party was not necessarily competitive in plurality contests, nor did it have a

---

<sup>34</sup> Some evidences imply that no party had recognized the effect of double candidacy system to competitiveness of plurality campaign. First, the drafters believe that the purpose of the double candidacy system is to keep small party candidates in the district contests, by allowing the PR candidates also to run plurality district (Kage 1997, pp313). Next, not only LDP, but also New Frontier Party (NFP), which achieved the second largest number of the seats in the first election after the reform, and one of the main actors in the electoral reform, did not use the double candidacy system to save their candidates lost in the plurality races. Moreover, Norihiko Narita, the former secretary of Prime Minister Hosokawa and one of the most important persons in the electoral reform, does not seem to believe that the double candidacy system was the important topic. He has lectured about the process more than 80 pages long (Narita 1996). In the lecture, he described detailed negotiation process for the electoral reform behind the scene. He argues that the main subject of the negotiation was the number of seat allocation to the plurality and PR tiers as well unit of the PR districts, but he did not mentioned anything about the double candidacy system in the hours long lecture.

sufficient number of incumbents or candidates to run in all plurality districts. Hence, originally the DPJ gave more emphasis to campaigning in the PR tier, not the plurality tier. In a sense, the LDP ran plurality candidates on the PR list using the double candidacy system, but the DPJ nominated the best PR candidates to plurality contests with the system (Kage 1997). In brief, the LDP and the DPJ chose different usages of the PR tier to solve their specific problems. Their decision had considerable impact on the fates of these two parties.

The magnitude of the incumbency advantage provides some implications for the current nature of electoral campaigns in Japan. There is a controversy over interpretation of the size of the incumbency advantage, because in some countries the advantage is even negative (Ariga 2010; Linden 2004; Titunik 2008; Uppal 2009). However, here I take a simple interpretation: the electorates in Japan are more likely to vote for the incumbent candidates and/or the candidates of the party holding incumbency, based on their name recognition, personal connection with the candidates, their access to government resources, and so on. The size, 2-3%, is not huge but not ignorable especially at competitive districts.

It is not clear yet that how the magnitude has changed between the electoral reform and today. Reed, Scheiner and Thies (2009) argue that the party affiliation has become more important factor than the candidate quality in the 2009 HR election. This suggests the transformation of electoral campaigns in Japan from candidate-centered to more administration and/or policy oriented. Nevertheless, it is still possible that the 2009 HR election was rather exceptional. The nature of electoral competition in Japan and its transformation should be examined further by carefully observing the subsequent elections.

Finally, this chapter demonstrates the importance of a seemingly minor electoral rule not only in the estimation of the variable of interest but also in its potential impact on democratic responsiveness of the electoral system. The country specific electoral rules like the double candidacy system are often neglected in cross-national studies. However, the result of this chapter indicates that these factors can have significant influence as

much as major and comparable features of the electoral system, such as the seat number balance between plurality and PR tiers of the mixed-member electoral system. This chapter shows an example of new direction for future research, which focuses on a single country but its implication is applicable to several countries.

## REFERENCE

- Angrist, Joshua D. and Victor C. Lavy. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics*. 114 (2): 533-575
- Ansolabehere, Stephen and James M. Snyder. 2002. "The Incumbency Advantage in U.S. Elections: An Analysis of State and Federal Offices, 1942–2000." *Election Law Journal: Rules, Politics and Policy*. 1 (3): 315-338.
- Ansolabehere, Stephen and James M. Snyder. 2004. "Using Term Limits to Estimate Incumbency Advantages When Officeholders Retire Strategically." *Legislative Studies Quarterly*. 29(4): 487–515.
- Ariga, Kenichi. 2010. "Entrenched Incumbents, Irresponsible Parties? Comparative Analysis of Incumbency Advantage across Different Electoral Systems." Dissertation manuscript
- Carson, Jamie L., Erik J. Engstrom, and Jason M. Roberts. 2007. "Candidate Quality, the Personal Vote, and the Incumbency Advantage in Congress." *American Political Science Review*. 101 (2): 289-301.
- Chay, Kenneth Y. and Michael Greenstone. 2005. "Does Air Quality Matter? Evidence from the Housing Market." *Journal of Political Economy*. 113 (2): 376-424.
- Christensen, Ray. 1998. "The Effect of Electoral Reforms on Campaign Practices in Japan: Putting New Wine into Old Bottles." *Asian Survey* 38 (10): 986-1004.
- Cox, Gary. 1997. *Making Votes Count: Strategic Coordination in the World's Electoral System*. Cambridge: Cambridge University Press.
- Cox, Gary and Jonathan N. Katz. 2002. *Elbridge Gerry's Salamander: The Electoral Consequences of the Reapportionment Revolution*. Cambridge: Cambridge University Press
- Curtis, Gerald L. 1971. *Election Campaigning, Japanese Style*. New York: Columbia University Press.
- Duverger, Maurice. 1954. *Political Parties, Their Organization and Activity in the Modern State*. New York, Wiley.
- Erikson, Robert S. 1971. "The Advantage of Incumbency in Congressional Elections." *Polity* 3 (3): 395-405.

- Fiorina, Morris. 1977. "The Case of the Vanishing Marginals: The Bureaucracy Did It." *American Political Science Review* 71 (1): 177-181.
- Gaines, Brian J. 1998. "The Impersonal Vote? Constituency Service and Incumbency Advantage in British Elections, 1950-1992." *Legislative Studies Quarterly* 23 (2): 167-195.
- Gelman, Andrew and Gary King. 1990. "Estimating Incumbency Advantage without Bias." *American Journal of Political Science* 34 (4): 1142-1164
- Griffin, John D. 2006. "Electoral Competition and Democratic Responsiveness: A Defense of the Marginality Hypothesis." *The Journal of Politics* 68(4): 911-921.
- Hahn, Jinyong, Petra Todd, and Wilbert H. Van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica* 69 (1): 201-209.
- Hainmueller, Jens and Holger L. Kern. 2008. "Incumbency as a Source of Contamination in Mixed Electoral Systems", *Electoral Studies* 27 (2): 213-227
- Hirano, Shigeo and James M. Snyder, Jr. 2009. "Using Multimember District Elections to Estimate the Sources of the Incumbency Advantage." *American Journal of Political Science* 53 (2): 292-306.
- Imbens, Guido and Joshua Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica*. 62 (2): 467-476.
- Kage, Rieko. 1997. "Institution Recognition and Realignment of Party System (seido ninshiki to seito sisutemu saihei)." in *The Studies on Political Realignment* (Seikai saihei no kenkyu : shin senkyo seido ni yoru sosankyō) by Otake Hideo ed. Tokyo : Yuhikaku. 303-338.
- Katz, Jonathan N., and Gary King. 1999. "A Statistical Model for Multiparty Electoral Data." *American Political Science Review* 93 (1): 15-32.
- King, Gary and Langche Zeng. 2007. "When Can History Be Our Guide? The Pitfalls of Counterfactual Inference". *International Studies Quarterly*. 51 (1): 183-210.
- Kohno, Masaru. 1997. *Japan's Postwar Party Politics*. Princeton, N.J.: Princeton University Press.



- Lee, David S. 2008. "Randomized Experiments from Non-Random Selection in US House Elections." *Journal of Econometrics* 142 (2): 675-697.
- Levitt, Steven D. and Catherine D Wolfram. 1997. "Decomposing the Sources of Incumbency Advantage in the US House". *Legislative Studies Quarterly* 22(1): 45-60.
- Linden, Leigh L. 2004. "Are Incumbents Really Advantaged? The Preference for Non-Incumbents in Indian National Elections." Unpublished manuscript.
- Mayhew, David R. 1974. "Congressional Elections: The Case of the Vanishing Marginals." *Polity* 6 (3): 295-317.
- McKean, Margaret and Ethan Scheiner. 2000. "Japan's New Electoral System: la plus ça change..." *Electoral Studies* 19 (4): 447-477.
- Narita, Norihiko. 1996. "How Prime Minister's Office and Coalition Parties Behaved for Political Reform Legislation - Witness of Chief Assistant to Prime Minister Hosokawa". *Hokudai Hogaku Ronsyu*: 46 (6) 406-486.
- Nichols, Austin. 2007. "RD: Stata Module for Regression Discontinuity Estimation." Boston College Department of Economics
- Niwa, Isao. 1997. "Activity of the LDP Local Organization- from a Case Study on Those in Toyama Prefecture (ziminto chiho sosiki no katsudo-toyama ken wo jirei to shite)." in *The Studies on Political Realignment* (Seikai saihen no kenkyu : shin senkyo seido ni yoru sosenkyo) by Otake Hideo ed. Tokyo : Yuhikaku. 253-300
- Porter, Jack. 2003. "Estimation in the Regression Discontinuity Model." Unpublished Manuscript.
- Powell, G. Bingham Jr. 2000. *Elections as Instruments of Democracy: Majoritarian and Proportional Visions*. New Haven: Yale University Press.
- Ramseyer, J. Mark, and Frances McCall Rosenbluth. 1993. *Japan's Political Marketplace*. Cambridge, Mass.: Harvard University Press.
- Scheiner, Ethan. 2005. *Democracy without Competition in Japan: Opposition Failure in a One-Party Dominant State*. Cambridge: Cambridge University Press.
- Taniguchi, Masaki. 2004. *Gendai Nihon no Senkyo Seiji [Electoral Reform in Japan]*. Tokyo: University of Tokyo Press.

- Titunik, Rocío. 2008. "Incumbency Advantage in Brazil: Evidence from Municipal Mayor Elections." University of California, Berkeley. Manuscript.
- Uppal, Yogesh. 2009. "The Disadvantaged Incumbents: Estimating Incumbency Effects in Indian State Legislatures." *Public Choice* 138 (1): 9-27.
- Van der Klaauw, Wilbert H. 2002. "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity." *International Economic Review* 43(4): 1249-1287.
- Wildavsky, Aaron. 1959. "A Methodological Critique of Duverger's Political Parties A Methodological Critique of Duverger's Political Parties." *The Journal of Politics* 21 (May): 303-318

## TABLES

Table 4.1: Incumbency Advantage with RDD

	<i>Bandwidth</i>	<i>Coefficient</i>	<i>Std.Error</i>
<i>RD: Incumbency Advantage</i>	.0250	-.008	(.025)
<i>RD: Incumbency Advantage</i>	.0125	-.005	(.033)
<i>RD: Incumbency Advantage</i>	.0500	-.011	(.015)

Table 4.2: Incumbency Advantage with Natural Experiment

	DPJ 2005	LDP led coalition 2005
<i>Zombie Effect</i>	-.021 * (.0110)	.034 * (.0173)
<i>LDP vote-share 2003</i>	4.233 (2.520)	2.780 (7.046)
<i>LDP vote-share 2003<sup>2</sup></i>	-28.846 (-18.73)	33.624 (137.7)
<i>SDPJ participates</i>	0.890 * (.0344)	N.A.
<i>SDPJ exits</i>	-0.032 * (.0146)	0.105 (.2470)
<i>Constant</i>	-1.656 (1.339)	-0.832 (3.445)
<i>Number of cases</i>	43	26
<i>Adjusted R<sup>2</sup></i>	.287	.254
<i>LDP vote-share range 2003</i>	[.538 .598]	[.460 .489]
<i>Vote-Share Ratio to Winner (Sekihai-Ritsu) 2003</i>	[.67 .86]	[.85 .95]

\*  $p < .05$  (single sided). Numbers in the parenthesis indicate the standard error of the estimates.

## Chapter 5

### Conclusion

Around twenty years have passed since many countries adopted the mixed-member electoral system in the later 1980s and the early 1990s, which include both the countries with long history of democracy such as Japan, Italy and New Zealand and those newly democratized in the East and the Central Europe. Nevertheless, the mechanisms of the system have not been well understood. Some advocates of the system expect that the system represents the best of single-member plurality and PR systems, because the system allows both party-centered campaign of the PR system and candidate responsiveness of the plurality system (Shugart and Wattenberg 2001). However, the political elites and/or the people might not to be so satisfied with the system. Among the countries that adopted the MMM system, Italy, Russia and Ukraine switched to the pure PR system in 2005. Moreover, New Zealand has a referendum in November 2011 if the country should keep the current MMP system<sup>35</sup>.

This dissertation seeks to clarify the mechanisms of the MMM system based on the Japanese case, which adopted the system to the HR in 1994 and conducted five elections under the system in 1996, 2000, 2003, 2005 and 2009. Chapters 2 and 3 focus on the strategic behavior of the small parties under the electoral system, by focusing the logic behind their nomination in the plurality electoral tier. In the Chapter 2, I reexamine the *contamination effect*: it is rational for small parties to nominate in plurality districts of the mixed-member system because they help to mobilize more PR votes. I demonstrate that two small leftist parties in Japan, the SDPJ and the JCP succeeded in mobilizing additional PR votes in the

---

<sup>35</sup> *New Zealand Herald* "MMP referendum to be held with 2011 election" Oct 20, 2009. [http://www.nzherald.co.nz/nz/news/article.cfm?c\\_id=1&objectid=10604317](http://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=10604317)

districts that they ran candidates: the *contamination effect* clearly exists. However, given the discretion of the parties in choosing the districts in which to nominate, they could have nominated candidates to the districts where they can achieve a larger gain than the other districts: the *ATT* should be larger than the *ATE*, although the previous studies implicitly assume that they are equal. Therefore, the average effect size would have been smaller if they had increased the number of districts. Hence, the (potential) impact of the contamination effect has been small at the national level.

In Chapter 3, I discuss an alternative purpose of the nominations: blackmailing major parties to choose policies more desirable to the small parties. I develop a formal model of the interactions between a small leftist party and a center-left major party over the policy position of the major party candidates. The model shows that the major party selects leftist rather than centrist candidates in a repeated game regardless of the median voters' preference for a centrist, if there is a small chance that the small leftist party is *ideological*. *Ideological* parties prefer the defeat of the centrist candidates of the major party to the victory of them, even if it could benefit their ideological enemy. I test the validity of the model by using the small leftist SDPJ's nomination in plurality contests and the major center-leftist DPJ's candidate positions in 2003 and 2005 HR elections. The empirical data clearly supports the model. The SDPJ is more likely to nominate if the DPJ nominate candidates with more rightist preference. In turn, the DPJ chooses the leftists when they pick up new candidates, if the SDPJ is popular in the districts.

In Chapter 4, I examine the magnitude of incumbency advantage in Japan after the electoral reform with a *natural experiment*. In the chapter, I utilize the mechanisms of the double candidacy system, which provide different incumbency status to the districts with very similar attributes but belonging to different PR regions. The model shows that the parties/candidates with PR incumbency performed 2-3% better in the plurality vote-share in the districts than those without seats, when they confronted with the incumbent opposition. The result shows that the incumbents in the Japanese HR still have some (but not huge) electoral advantage owing to their office. On the other hand, the result implies that the double candidacy system hurts the district incumbents, counter to the conventional wisdom. The system awards available PR seats to the most competitive candidates in each district,

who failed to win at the district level. These MPs challenge at the district level with PR incumbency status from the PR tier, which cancels out the advantage of the district incumbents.

I believe that these works demonstrate examples how the study on a single country, Japan, could contribute to the development of the theory in comparative politics. Japanese politics has not been well studied in a comparative perspective with few great exceptions, owing to its geographical and language distance from the other traditional democratic countries. The discussion in Chapter 2, especially that regarding the difference between the *ATT* and the *ATE*, is applicable not only to the contamination effect in the other mixed-member electoral system, but also any other political studies in which the treatment assignment is decided by the deliberation of the actors. The implication of the model in Chapter 3 is relevant to negotiations in any plurality electoral systems, not just the plurality tier of the mixed-member electoral system. The result of Chapter 4 shows that the double candidacy system, seemingly minor electoral rule and neglected in comparative studies, has significant impact to the incumbency advantage and the competitiveness of plurality contests.

Moreover, these studies demonstrate directions of the future in political studies. Chapter 2 utilizes the theoretical development of causal inference in statistics to refine the theoretical concepts to be measured. Chapter 4 uses a natural experiment to estimate the treatment effect with less relying on parametric assumptions of the model. These works indicate examples of the studies designed to achieve more reliable inference from observational data, which is currently dominant in political science. In turn, Chapter 3 shows the integration of the theoretical models and empirical evaluation. In combination with substantive knowledge regarding political phenomena, the development of these methodological approaches will provide fruitful results in the field.

## REFERENCE

Shugart, Matthew S., and Martin P. Wattenberg ed. 2001. *Mixed-member Electoral Systems: the Best of Both Worlds?* Oxford, UK: Oxford University Press.